

From: Chan, Christina
Sent: Thursday, May 08, 2003 3:00 PM
To: Yu, Misook; STIC-Biotech/ChemLib
Subject: RE: 09/851422 == erroneous file processed

Please rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

-----Original Message-----

From: Yu, Misook
Sent: Thursday, May 08, 2003 2:56 PM
To: Chan, Christina
Subject: FW: 09/851422 == erroneous file processed

The case is 8 month amended and previous search was improperly done. Please approve rush search for SEQ ID NO:1, 2, and 8.

Examiner Misook Yu, Ph.D.
703-308-2454 (Phone)
Art Unit 1642
CM1-8E18 (Room)
CM1-8E12 (Mail Box)

-----Original Message-----

From: Corrigan, Anne-Marie
Sent: Thursday, May 08, 2003 1:40 PM
To: Yu, Misook
Cc: Mosher, Mary
Subject: 09/851422 == erroneous file processed

Exr. Yu:

Unfortunately, STIC processed the wrong file under the above serial number (listed in search system as US09851422B). I removed the erroneous files from the search system; now, the files corresponding to the "ENTERED" raw sequence listing (09/851422A) in your file wrapper can be searched. If you'd like, I'd be happy to contact the searcher who produced your search.

I apologize for the error.

RECEIVED
MAY - 8 2003
STIC/CHEN, L. (STIC)

Point of Contact:
Thomas G. Larson, Ph.D.
703-308-7309
CM1, Rm. 6 B 01

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 5/8
Date Completed: 5/12
Searcher Prep/Review: 15
Clerical: _____
Online time: 15

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: 3
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:07:45 ; Search time 52.2353 Seconds
(without alignments)
94.386 Million cell updates/sec

Title: US-09-851-422A-1
Perfect score: 138
Sequence: 1 GFIAFLCTKXXXXVLDKFKXXXXXLIQIEDKXXXX 37

Scoring table: BIOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq_101002:*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	73.9	25	AAO21730	Amoebapore helix 3
2	102	73.9	25	AAO21730	Procytotoxin cytol
3	102	73.9	25	AAE18192	Entamoeba histolyt
4	102	73.9	25	AAE18197	Procytotoxin #1 fo
5	102	73.9	26	AAO21737	Procytotoxin cytol
6	102	73.9	26	AAO21738	Procytotoxin cytol
7	102	73.9	26	AAE18200	Procytotoxin #1 for tr
8	102	73.9	27	AAO21736	Procytotoxin cytol
9	102	73.9	27	AAE18199	Procytotoxin #3 fo
10	101.5	73.6	27	AAE18201	Procytotoxin #2 for tr

11	97.5	70.7	24	AAO21741	Procytotoxin cytol
12	97.5	70.7	24	AAE18203	Cytolytic peptide
13	97.5	70.7	26	AAE18204	Procytolytic pepti
14	97.5	70.7	26	AAE18205	Procytotoxin #5 fo
15	44	31.9	463	AAE1753	Arabidopsis thaila
16	44	31.9	493	AAE1752	Arabidopsis thaila
17	44	31.9	600	AAE1751	Arabidopsis thaila
18	43	31.2	431	AAE1750	Human secreted pro
19	43	31.2	617	AAE1749	Novel human diagno
20	43	31.2	626	AAE1748	Novel human diagno
21	42	30.4	893	AAE1747	The DNA polymerase
22	42	30.4	893	AAE1746	The DNA polymerase
23	42	30.4	893	AAE1745	The DNA polymerase
24	42	30.4	893	AAE1744	The DNA polymerase
25	42	30.4	893	AAE1743	The DNA polymerase
26	42	30.4	893	AAE1742	The DNA polymerase
27	42	30.4	893	AAE1741	The DNA polymerase
28	42	30.4	893	AAE1740	The DNA polymerase
29	42	30.4	893	AAE1739	The DNA polymerase
30	42	30.4	893	AAE1738	The DNA polymerase
31	42	30.4	893	AAE1737	The DNA polymerase
32	42	30.4	893	AAE1736	The DNA polymerase
33	42	30.4	893	AAE1735	The DNA polymerase
34	42	30.4	893	AAE1734	The DNA polymerase
35	42	30.4	893	AAE1733	The DNA polymerase
36	42	30.4	893	AAE1732	The DNA polymerase
37	42	30.4	893	AAE1731	The DNA polymerase
38	42	30.4	893	AAE1730	The DNA polymerase
39	42	30.4	893	AAE1729	The DNA polymerase
40	42	30.4	893	AAE1728	The DNA polymerase
41	42	30.4	893	AAE1727	The DNA polymerase
42	42	30.4	893	AAE1726	The DNA polymerase
43	42	30.4	893	AAE1725	The DNA polymerase
44	42	30.4	893	AAE1724	The DNA polymerase
45	42	30.4	893	AAE1723	The DNA polymerase

ALIGNMENTS

RESULT 1	AAO21730	standard; Peptide: 25 AA.
ID	AAO21730	
XX	AAO21730	
AC	AAO21730	
XX	13-SEP-2002 (first entry)	
DT	13-SEP-2002 (first entry)	
XX	Amoebapore helix 3 cytolitic peptide.	
DE	Amoebapore helix 3 cytolitic peptide.	
XX	Cytotoxic; cytosolic; procytotoxin; inactivator; protease; cancer;	
KW	ovary; prostate; breast; skin; lung; pancreas; Amoebapore helix 3.	
XX	Entamoeba histolytica.	
OS	Entamoeba histolytica.	
XX	US2002045736-A1.	
PN	US2002045736-A1.	
XX	18-APR-2002.	
PD	18-APR-2002.	
XX	27-AUG-2001; 2001US-0938623.	
PF	27-AUG-2001; 2001US-0938623.	
XX	09-MAY-2001; 2001US-0851422.	
PR	09-MAY-2001; 2001US-0851422.	
XX	(YUXX/) YU X.	
PA	(WAGN/) WAGNER T E.	
XX	Yu X. Wagner TE;	
PI	Yu X. Wagner TE;	
XX	WPI; 2002-507251/54.	
DR	WPI; 2002-507251/54.	
XX	A new procytotoxin useful in the treatment of cancer of e.g. prostate,	
PT	ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator	
XX	via a peptide bond cleavable by a specific protease	

XX Disclosure: Page 3; 21pp; English.
 PS The invention relates to a procytotoxin comprising a cytotoxic peptide
 CC bound to an inactivator via a peptide bond, where the peptide bond is
 CC susceptible to cleavage by a targeting specific protease. The
 CC procytotoxin is used to treat cancer, particularly of the prostate,
 CC ovary, breast, skin, lung or pancreas. This sequence represents the
 CC Amoebapore helix 3 cytolytic peptide relating to the invention.
 XX
 SO Sequence 25 AA:
 Query Match 73.9%; Score 102; DB 23; Length 25;
 Best Local Similarity 75.8%; Pred. No. 5.5e-12;
 Matches 25; Conservative 0; Mismatches 0; Indels 8; Gaps 2;
 Oy 1 GFATLCTKXXXXVDFGIDKXXXXLIQLIEDK 33
 Db 1 GFATLCTK----VDFGIDK----LIQLIEDK 25
 RESULT 2
 AAO21739
 ID AAO21739 standard; Peptide: 25 AA.
 AC AAO21739;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Procycytotoxin cytolytic peptide #4.
 XX
 KW Cytotoxic; cytostatic; procycytotoxin; inactivator; protease; cancer;
 KW ovary; prostate; breast; skin; lung; pancreas; cytolytic.
 XX
 OS Unidentified.
 FH
 XX Key Location/Qualifiers
 FT Modified-site 9
 FT /note="This Lys residue is modified by R, where R is
 FT independently selected from the group consisting of the
 FT epsilon-amino group of the adjacent lysine residue,
 FT (epsilon-gamma)-Glu, (epsilon-gamma)-Glu-(alpha-gamma)-
 FT (Glu)-1-3, (epsilon-alpha)-(Phe)-1-3, (epsilon-alpha)-
 FT (Tyr)-1-3, (epsilon-alpha)-(Trp)-1-3, (epsilon-alpha)-
 FT (Lys)-1-3, (epsilon-alpha)-(Arg)-1-3, wherein (epsilon-
 FT gamma) represents a peptide bond between the epsilon
 FT amino group of lysine and the gamma carboxyl group of the
 FT adjacent glutamate, (alpha-gamma) represents a peptide
 FT bond between the alpha amino group of the first glutamate
 FT and the gamma carboxyl group of the second glutamate,
 FT (epsilon-alpha) represents a peptide bond between the
 FT epsilon-amino acid of lysine and the alpha carboxyl group
 FT of the indicated amino acid"
 FT 17
 FT /note="This Lys residue is modified by R, where R is
 FT independently selected from the group consisting of the
 FT epsilon-amino group of the adjacent lysine residue,
 FT (epsilon-gamma)-Glu, (epsilon-gamma)-Glu-(alpha-gamma)-
 FT (Glu)-1-3, (epsilon-alpha)-(Phe)-1-3, (epsilon-alpha)-
 FT (Tyr)-1-3, (epsilon-alpha)-(Trp)-1-3, (epsilon-alpha)-
 FT (Lys)-1-3, (epsilon-alpha)-(Arg)-1-3, wherein (epsilon-
 FT gamma) represents a peptide bond between the epsilon
 FT amino group of lysine and the gamma carboxyl group of the
 FT adjacent glutamate, (alpha-gamma) represents a peptide
 FT bond between the alpha amino group of the first glutamate
 FT and the gamma carboxyl group of the second glutamate,
 FT (epsilon-alpha) represents a peptide bond between the
 FT epsilon-amino acid of lysine and the alpha carboxyl group
 FT of the indicated amino acid"
 FT 25
 FT /note="This Lys residue is modified by R, where R is
 FT independently selected from the group consisting of the
 FT epsilon-amino group of the adjacent lysine residue,
 FT (epsilon-gamma)-Glu, (epsilon-gamma)-Glu-(alpha-gamma)-
 FT (Glu)-1-3, (epsilon-alpha)-(Phe)-1-3, (epsilon-alpha)-
 FT (Tyr)-1-3, (epsilon-alpha)-(Trp)-1-3, (epsilon-alpha)-
 FT (Lys)-1-3, (epsilon-alpha)-(Arg)-1-3, wherein (epsilon-
 FT gamma) represents a peptide bond between the epsilon
 FT amino group of lysine and the gamma carboxyl group of the
 FT adjacent glutamate, (alpha-gamma) represents a peptide
 FT bond between the alpha amino group of the first glutamate
 FT and the gamma carboxyl group of the second glutamate,
 FT (epsilon-alpha) represents a peptide bond between the
 FT epsilon-amino acid of lysine and the alpha carboxyl group
 FT of the indicated amino acid"
 FT 25
 FT Modified-site
 FT 25
 FT /note="This Lys residue is modified by R, where R is
 FT independently selected from the group consisting of the
 FT epsilon-amino group of the adjacent lysine residue,
 FT (epsilon-gamma)-Glu, (epsilon-gamma)-Glu-(alpha-gamma)-
 FT (Glu)-1-3, (epsilon-alpha)-(Phe)-1-3, (epsilon-alpha)-
 FT (Tyr)-1-3, (epsilon-alpha)-(Trp)-1-3, (epsilon-alpha)-
 FT (Lys)-1-3, (epsilon-alpha)-(Arg)-1-3, wherein (epsilon-
 FT gamma) represents a peptide bond between the epsilon
 FT amino group of lysine and the gamma carboxyl group of the
 FT adjacent glutamate, (alpha-gamma) represents a peptide
 FT bond between the alpha amino group of the first glutamate
 FT and the gamma carboxyl group of the second glutamate,
 FT (epsilon-alpha) represents a peptide bond between the
 FT epsilon-amino acid of lysine and the alpha carboxyl group
 FT of the indicated amino acid"
 FT 25
 FT Modified-site
 FT 25
 FT /note="This Lys residue is modified by R, where R is
 FT independently selected from the group consisting of the
 FT epsilon-amino group of the adjacent lysine residue,
 FT (epsilon-gamma)-Glu, (epsilon-gamma)-Glu-(alpha-gamma)-
 FT (Glu)-1-3, (epsilon-alpha)-(Phe)-1-3, (epsilon-alpha)-
 FT (Tyr)-1-3, (epsilon-alpha)-(Trp)-1-3, (epsilon-alpha)-
 FT (Lys)-1-3, (epsilon-alpha)-(Arg)-1-3, wherein (epsilon-
 FT gamma) represents a peptide bond between the epsilon
 FT amino group of lysine and the gamma carboxyl group of the
 FT adjacent glutamate, (alpha-gamma) represents a peptide
 FT bond between the alpha amino group of the first glutamate
 FT and the gamma carboxyl group of the second glutamate,
 FT (epsilon-alpha) represents a peptide bond between the
 FT epsilon-amino acid of lysine and the alpha carboxyl group
 FT of the indicated amino acid"

FT (epsilon-gamma)-Glu, (epsilon-gamma)-Glu-(alpha-gamma)-
 FT (Glu)-1-3, (epsilon-alpha)-(Phe)-1-3, (epsilon-alpha)-
 FT (Tyr)-1-3, (epsilon-alpha)-(Trp)-1-3, (epsilon-alpha)-
 FT (Lys)-1-3, (epsilon-alpha)-(Arg)-1-3, wherein (epsilon-
 FT gamma) represents a peptide bond between the epsilon
 FT amino group of lysine and the gamma carboxyl group of the
 FT adjacent glutamate, (alpha-gamma) represents a peptide
 FT bond between the alpha amino group of the first glutamate
 FT and the gamma carboxyl group of the second glutamate,
 FT (epsilon-alpha) represents a peptide bond between the
 FT epsilon-amino acid of lysine and the alpha carboxyl group
 FT of the indicated amino acid"
 FT 25
 FT US2002045736-A1.
 PN 18-APR-2002.
 PD 27-AUG-2001; 2001US-0938623.
 XX 09-MAY-2001; 2001US-0851422.
 PR (YUXX/) YU X.
 XX (MAGN/) MAGNER T E.
 PA (MAGN/) MAGNER T E.
 XX Yu X, Wagner TE;
 PI WPI: 2002-507251/54.
 DR
 XX A new procycytotoxin useful in the treatment of cancer of e.g. prostate,
 PT ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator
 PT via a peptide bond cleavable by a specific protease
 PT
 PS Disclosure: Page 7; 21pp; English.
 XX
 CC The invention relates to a procycytotoxin comprising a cytotoxic peptide
 CC bound to an inactivator via a peptide bond, where the peptide bond is
 CC susceptible to cleavage by a targeting specific protease. The
 CC procycytotoxin is used to treat cancer, particularly of the prostate,
 CC ovary, breast, skin, lung or pancreas. This sequence represents a
 CC procycytotoxin cytolytic peptide sequence relating to the invention.
 XX
 SQ Sequence 25 AA:
 Query Match 73.9%; Score 102; DB 23; Length 25;
 Best Local Similarity 75.8%; Pred. No. 5.5e-12;
 Matches 25; Conservative 0; Mismatches 0; Indels 8; Gaps 2;
 Oy 1 GFATLCTKXXXXVDFGIDKXXXXLIQLIEDK 33
 Db 1 GFATLCTK----VDFGIDK----LIQLIEDK 25
 RESULT 3
 AAE18192
 ID AAE18192 standard; peptide: 25 AA.
 AC AAE18192;
 XX
 DT 07-MAY-2002 (first entry)
 XX
 DE Entamoeba histolytica procycytotoxin, amoebapore helix 3.
 XX
 KW Procycytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
 KW prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
 XX
 OS Entamoeba histolytica.
 XX
 PN WO200185777-A2.
 PD 15-NOV-2001.
 PD 09-MAY-2001; 2001WO-US40690.
 XX

PR	09-MAY-2000: 2000US-203063P.
PR	16-JUN-2000: 2000US-212042P.
XX	
PA	(GREE-) GREENVILLE HOSPITAL SYSTEM.
XX	
PI	Yu X, Magnier TE;
XX	
DR	WPI: 2002-154423/20.
XX	
PT	A procytotoxin comprising a cytotoxic peptide with at least one lysine
PT	residue which is converted into a cytotoxin in a target cell is used
PT	for treating cancer, particularly of the prostate, skin, ovary or lung
XX	
XX	Disclosure; Page 7; 33pp; English.
XX	
CC	The present invention relates to a procytotoxin, comprising a cytotoxic
CC	peptide with at least one lysine residue bound by a peptide bond to at
CC	least one amino acid via the epsilon-amino group of the lysine residue.
CC	The procytotoxin is used to treat cancer, particularly of the prostate,
CC	ovary, lung or skin. Cytolytic peptide also known as channel-forming
CC	peptide, typically disrupt cell membrane, causing cell lysis and death
CC	(apoptosis) upon contact. The present sequence is Entamoeba histolytica
CC	procytotoxin, amoebapore helix 3.
XX	
SO	Sequence 25 AA:
	Query Match 73.9%; Score 102; DB 23; Length 25;
	Best Local Similarity 75.8%; Pred. No. 5.5e-12;
	Matches 25; Conservative 0; Mismatches 0; Indels 8; Gaps 2.
Oy	1 GFIAFLCRKXXXVLDFGIDKXXXLIQLIEDK 33
Db	1 GFIAFLCRK---VLDFGIDK---LIQLIEDK 25
RESULT 4	
AAE18197	
ID	AAE18197 standard; peptide; 25 AA.
XX	
AC	AAE18197;
XX	
DT	07-MAY-2002 (first entry)
XX	
DE	Procytotoxin #1 for treating cancer.
XX	
KM	Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
KW	prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX	
OS	Unidentified.
XX	
Key	
FH	Modified-site
FT	
FT	9 location/Qualifiers
FT	/note- "Lys(R), where R is the unmodified epsilon-amino
FT	group of the adjacent Lys, or is [epsilon]-gamma)-Glu,
FT	[epsilon]-gamma)-Glu-(alpha-gamma)-(Glu)1-3,
FT	[epsilon]-alpha)-(Phe)1-3, [epsilon]-alpha)-(Tyr)1-3,
FT	[epsilon]-alpha)-(Trp)1-3, [epsilon]-alpha)-(Lys)1-3 or
FT	[epsilon]-alpha)-(Arg)1-3; where [epsilon]-gamma)
FT	represents a peptide bond between the epsilon amino
FT	group of lysine and the gamma carboxyl group of the
FT	adjacent glutamate, [alpha-gamma) represents a peptide
FT	bond between the alpha amino group of the first glutamate
FT	and the gamma carboxyl group of the second glutamate,
FT	[epsilon]-alpha) represents a peptide bond between the
FT	epsilon amino acid of lysine and the alpha carboxyl
FT	group of the indicated amino acid, and 1-3 indicates that
FT	1, 2 or 3 of the designated amino acid can be linked to
FT	the first via conventional peptide bonds"
FT	17
FT	/note- "Lys(R), where R is the unmodified epsilon-amino
FT	group of the adjacent Lys, or is [epsilon]-gamma)-Glu,
FT	[epsilon]-gamma)-Glu-(alpha-gamma)-(Glu)1-3,

FT [epsilon]ion-alpha)-(Phe)1-3; [epsilon]ion-alpha)-(Tyr)1-3,
FT [epsilon]ion-alpha)-(Trp)1-3; [epsilon]ion-alpha)-(Lys)1-3 or
FT [epsilon]ion-alpha)-(Arg)1-3; where [epsilon]ion-gamma)
FT represents a peptide bond between the epsilon amino
FT group of lysine and the gamma carboxyl group of the
FT adjacent glutamate, [alpha]-gamma) represents a peptide
FT bond between the alpha amino group of the first glutamate
FT and the gamma carboxyl group of the second glutamate,
FT [epsilon]ion-alpha) represents a peptide bond between the
FT epsilon amino acid of lysine and the alpha carboxyl
FT group of the indicated amino acid, and 1-3 indicates that
FT 1, 2 or 3 of the designated amino acid can be linked to
FT the first via conventional peptide bonds"
FT 25

Modified-site
FT /note- "Lys(R), where R is the unmodified epsilon-amino
FT group of the adjacent Lys, or is [epsilon]ion-gamma)-Glu,
FT [epsilon]ion-gamma)-Glu-(alpha-gamma)-(Glu)1-3,
FT [epsilon]ion-gamma)-(Phe)1-3; [epsilon]ion-alpha)-(Tyr)1-3,
FT [epsilon]ion-alpha)-(Trp)1-3; [epsilon]ion-alpha)-(Lys)1-3 or
FT [epsilon]ion-alpha)-(Arg)1-3; where [epsilon]ion-gamma)
FT represents a peptide bond between the epsilon amino
FT group of lysine and the gamma carboxyl group of the
FT adjacent glutamate, [alpha]-gamma) represents a peptide
FT bond between the alpha amino group of the first glutamate
FT and the gamma carboxyl group of the second glutamate,
FT [epsilon]ion-alpha) represents a peptide bond between the
FT epsilon amino acid of lysine and the alpha carboxyl
FT group of the indicated amino acid, and 1-3 indicates that
FT 1, 2 or 3 of the designated amino acid can be linked to
FT the first via conventional peptide bonds"

WO200185777-A2.
XX PD
XX 15-NOV-2001.
XX PF 09-MAY-2001; 2001WO-US40690.
XX PR 09-MAY-2000; 2000US-203063P.
XX PR 16-JUN-2000; 2000US-212042P.
XX PA (GREE-) GREENVILLE HOSPITAL SYSTEM.

XX PI Yu X, Wagner TE;
XX DR WPI; 2002-154423/20.

PT A procytotoxin comprising a cytotoxic peptide with at least one lysine
PT residue which is converted into a cytoloxin in a target cell is used
PT for treating cancer, particularly of the prostate, skin, ovary or lung
PT -
XX PS
XX Claim 5; Page 23; 33pp; English.

XX CC The present invention relates to a procytotoxin, comprising a cytotoxic
CC peptide with at least one lysine residue bound by a peptide bond to at
CC least one amino acid via the epsilon-amino group of the lysine residue.
CC The procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
CC peptide, typically disrupt cell membrane, causing cell lysis and death
CC (apoptosis) upon contact. The present sequence is procytotoxin.
XX SQ

Sequence 25 AA:
Query March 73.9%; Score 102; DB 23; Length 25;
Best Local Similarity 75.8%; Pred. No. 5.5e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 8; Gaps 2;

OY 1 GFIAFLCTKXXXXVLDFGIDKXXXXLIOLIEDK 33
| | | | | | | | | | | | | | | | | | | | | |
Db 1 GFIAFLCTK----VI DFGIDK-----LIQLIEDK 25

RESULT 5

AA021737
ID AA021737 standard; Peptide: 26 AA.
XX
AC AA021737;
XX
DT 13-SEP-2002 (first entry)
XX
DE procytotoxin cytolytic peptide #2.
XX
XX Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;
KW ovary; prostate; breast; skin; lung; pancreas; cytolytic.
XX
OS unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 25 /note= "The epsilon amino group of Lys 1s peptide bonded
FT to the alpha carboxyl group of Phe"
FT
FT Modified-site 26 /note= "The alpha carboxyl group of Phe is peptide
FT bonded to the epsilon amino group of Lys"
XX
XX US2002045736-A1.
XX
XX 18-APR-2002.
XX
XX 27-AUG-2001; 2001US-0938623.
XX
XX 09-MAY-2001; 2001US-0851422.
XX
XX (YUXX/) YU X.
XX (MAGN/) MAGNER T E.
XX
XX YU X, Wagner TE;
XX
XX WPI: 2002-507251/54.
XX
XX A new procytotoxin useful in the treatment of cancer of e.g. prostate,
PT ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator
PT via a peptide bond cleavable by a specific protease
XX
XX Disclosure: Page 7; 21pp; English.
XX
XX The invention relates to a procytotoxin comprising a cytotoxic peptide
CC bound to an inactivator via a peptide bond, where the peptide bond is
CC susceptible to cleavage by a targeting specific protease. The
CC procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, breast, skin, lung or pancreas. This sequence represents a
CC procytotoxin cytolytic peptide sequence relating to the invention.
XX
XX Sequence 26 AA:
SQ
Query Match 73.9%; Score 102; DB 23; Length 26;
Best Local Similarity 75.8%; Pred. No. 5.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 8; Gaps 2;
OY 1 GFIATLCTKXXXXVLDFGIDKXXXXLIQLIEDK 33
DB 1 GFIATLCTK----VLDFGIDK----LIQLIEDK 25
RESULT 6
AA021738
ID AA021738 standard; Peptide: 26 AA.
XX
AC AA021738;
XX
XX 13-SEP-2002 (first entry)
XX
XX Procytotoxin cytolytic peptide #3.
XX
XX Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;
KW ovary; prostate; breast; skin; lung; pancreas; cytolytic.

XX
OS unidentified.
XX
XX Key Location/Qualifiers
FH Modified-site 17 /note= "The epsilon amino group of Lys 1s peptide bonded
FT to the alpha carboxyl group of Phe"
FT
FT Modified-site 25 /note= "The epsilon amino group of Lys 1s peptide bonded
FT to the alpha carboxyl group of Phe"
FT
FT Modified-site 26 /note= "The alpha carboxyl group of Phe is peptide bonded
FT to the epsilon amino group of Lys"
XX
XX US2002045736-A1.
XX
XX 18-APR-2002.
XX
XX 27-AUG-2001; 2001US-0938623.
XX
XX 09-MAY-2001; 2001US-0851422.
XX
XX (YUXX/) YU X.
XX (MAGN/) MAGNER T E.
XX
XX YU X, Wagner TE;
XX
XX WPI: 2002-507251/54.
XX
XX A new procytotoxin useful in the treatment of cancer of e.g. prostate,
PT ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator
PT via a peptide bond cleavable by a specific protease
XX
XX Disclosure: Page 7; 21pp; English.
XX
XX The invention relates to a procytotoxin comprising a cytotoxic peptide
CC bound to an inactivator via a peptide bond, where the peptide bond is
CC susceptible to cleavage by a targeting specific protease. The
CC procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, breast, skin, lung or pancreas. This sequence represents a
CC procytotoxin cytolytic peptide sequence relating to the invention.
XX
XX Sequence 26 AA:
SQ
Query Match 73.9%; Score 102; DB 23; Length 26;
Best Local Similarity 75.8%; Pred. No. 5.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 8; Gaps 2;
OY 1 GFIATLCTKXXXXVLDFGIDKXXXXLIQLIEDK 33
DB 1 GFIATLCTK----VLDFGIDK----LIQLIEDK 25
RESULT 7
AAE18200
ID AAE18200 standard; peptide: 26 AA.
XX
AC AAE18200;
XX
XX 07-MAY-2002 (first entry)
XX
XX Procytin #1 for treating cancer.
XX
XX Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
KW prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX
XX unidentified.
XX
XX Key Location/Qualifiers
FT Modified-site 25 /note= "Lys-[epsilon-alpha]; where [epsilon-alpha]
FT represents a peptide bond between the epsilon amino
FT group of lysine and the alpha carboxyl group of the

FT adjacent phenylalanine"
 XX
 XX WO200185777-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 09-MAY-2001; 2001WO-US40690.
 XX
 PR 09-MAY-2000; 2000US-203063P.
 PR 16-JUN-2000; 2000US-212042P.
 XX
 PA (GREE-) GREENVILLE HOSPITAL SYSTEM.
 XX
 PI Yu X, Wagner TE;
 XX
 DR WPI: 2002-154423/20.
 XX
 XX A procytotoxin comprising a cytotoxic peptide with at least one lysine
 PT residue which is converted into a cytotoxin in a target cell is used
 PT for treating cancer, particularly of the prostate, skin, ovary or lung
 PT
 PS Disclosure: Page 12; 33pp; English.
 XX
 CC The present invention relates to a procytotoxin, comprising a cytotoxic
 CC peptide with at least one lysine residue bound by a peptide bond to at
 CC least one amino acid via the epsilon-amino group of the lysine residue.
 CC The procytotoxin is used to treat cancer, particularly of the prostate,
 CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
 CC peptide, typically disrupt cell membrane, causing cell lysis and death
 CC (apoptosis) upon contact. The present sequence is protoxin.
 CC
 SQ Sequence 26 AA;
 Query Match 73.9%; Score 102; DB 23; Length 26;
 Best Local Similarity 75.8%; Pred. No. 5.7e-12;
 Matches 25; Conservative 0; Mismatches 0; Indels 8; Gaps 2;
 OY 1 GFATLCTKXXXVLDPFGIDKXXXLIQLIEDK 33
 ||||||| ||||||| |||||||
 Db 1 GFATLCTK----VLDPFGIDK----LIQLIEDK 25
 RESULT 8
 AAO21736
 ID AAO21736 standard; Peptide: 27 AA.
 XX
 AC AAO21736;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Procytotoxin cytolytic peptide #1.
 XX
 KW Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;
 KW ovary; prostate; breast; skin; lung; pancreas; cytolytic.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 25
 /note- "The epsilon amino group of Lys is peptide bonded
 FT to the gamma carboxyl group of Glu"
 FT 26
 /note- "The gamma carboxyl group of this Glu is peptide
 FT bonded to the epsilon amino group of Lys. The alpha amino
 FT group of this Glu is peptide bonded to the gamma carboxyl
 FT group of another Glu"
 FT 27
 /note- "The gamma carboxyl group of this Glu is peptide
 FT bonded to the alpha amino group of another Glu"
 XX
 PN US2002045736-A1.
 XX

PD 18-APR-2002.
 XX
 PF 27-AUG-2001; 2001US-0938623.
 XX
 PR 09-MAY-2001; 2001US-0851422.
 XX
 PA (YUXX/) YU X.
 PA (WAGN/) WAGNER T E.
 XX
 PI Yu X, Wagner TE;
 XX
 DR WPI: 2002-507251/54.
 XX
 XX A new procytotoxin useful in the treatment of cancer of e.g. prostate,
 PT ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator
 PT via a peptide bond cleavable by a specific protease
 PT
 PS Disclosure: Page 7; 21pp; English.
 XX
 CC The invention relates to a procytotoxin comprising a cytotoxic peptide
 CC bound to an inactivator via a peptide bond, where the peptide bond is
 CC susceptible to cleavage by a targeting specific protease. The
 CC procytotoxin is used to treat cancer, particularly of the prostate,
 CC ovary, breast, skin, lung or pancreas. This sequence represents a
 CC procytotoxin cytolytic peptide sequence relating to the invention.
 CC
 SQ Sequence 27 AA;
 Query Match 73.9%; Score 102; DB 23; Length 27;
 Best Local Similarity 75.8%; Pred. No. 5.9e-12;
 Matches 25; Conservative 0; Mismatches 0; Indels 8; Gaps 2;
 OY 1 GFATLCTKXXXVLDPFGIDKXXXLIQLIEDK 33
 ||||||| ||||||| |||||||
 Db 1 GFATLCTK----VLDPFGIDK----LIQLIEDK 25
 RESULT 9
 AAE18199
 ID AAE18199 standard; peptide: 27 AA.
 XX
 AC AAE18199;
 XX
 DT 07-MAY-2002 (first entry)
 XX
 DE Procytotoxin #3 for treating cancer.
 XX
 KW Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
 KW prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 25
 /note- "Lys-[epsilon-gamma]; where [epsilon-gamma]
 FT represents a peptide bond between the epsilon amino
 FT group of lysine and the gamma carboxyl group of the
 FT adjacent glutamate"
 FT 26
 /note- "Glu-[alpha-gamma]; where [alpha-gamma]
 FT represents a peptide bond between the alpha
 FT amino group of the first glutamate and the gamma carboxyl
 FT group of the second glutamate"
 XX
 PN WO200185777-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 09-MAY-2001; 2001WO-US40690.
 XX
 PR 09-MAY-2000; 2000US-203063P.
 PR 16-JUN-2000; 2000US-212042P.
 XX

PA (Gree-) GREENVILLE HOSPITAL SYSTEM.
 XX
 PI Yu X, Wagner TE;
 XX
 DR WPI: 2002-154423/20.
 XX
 PT A procytotoxin comprising a cytotoxic peptide with at least one lysine
 residue which is converted into a cytotoxin in a target cell is used
 PT for treating cancer, particularly of the prostate, skin, ovary or lung
 PT
 XX
 PS Disclosure: Page 11; 33pp; English.
 XX
 CC The present invention relates to a procytotoxin, comprising a cytotoxic
 CC peptide with at least one lysine residue bound by a peptide bond to at
 CC least one amino acid via the epsilon-amino group of the lysine residue.
 CC The procytotoxin is used to treat cancer, particularly of the prostate,
 CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
 CC peptide, typically disrupt cell membrane, causing cell lysis and death
 CC (apoptosis) upon contact. The present sequence is procytotoxin.
 CC
 XX
 SQ Sequence 27 AA:
 Query Match 73.9%; Score 102; DB 23; Length 27;
 Best Local Similarity 75.8%; Pred. No. 5.9e-12;
 Matches 25; Conservative 0; Mismatches 0; Indels 8; Gaps 2;
 OY 1 GFATLCTKXXXXVLDGIDKXXXXLIQLIEDK 33
 ||||| ||||| ||||| |||||
 DB 1 GFATLCTK----VLDGIDK----LIQLIEDK 25
 RESULT 10
 AAE18201
 ID AAE18201 standard: peptide: 27 AA.
 XX
 AC AAE18201:
 XX
 DT 07-MAY-2002 (first entry)
 XX
 DE Proctoxin #2 for treating cancer.
 XX
 KW Procytotoxin: cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
 KW prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
 OS
 XX
 OS Unidentified.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 17..18
 FT /note= "Lys([epsilon]-alpha)-Phe): where ([epsilon]-alpha)
 FT represents a peptide bond between the epsilon amino
 FT group of lysine and the alpha carboxyl group of
 FT phenylalanine, and a standard peptide linkage between
 FT lysine and phenylalanine"
 FT 26
 FT /note= "Lys-[epsilon]-alpha): where ([epsilon]-alpha)
 FT represents a peptide bond between the epsilon amino
 FT group of lysine and the alpha carboxyl group of the
 FT adjacent phenylalanine"
 FT
 PN W0200185777-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 09-MAY-2001: 2001WO-US40690.
 XX
 PR 09-MAY-2000: 2000US-203063P.
 PR 16-JUN-2000: 2000US-212042P.
 XX
 PA (Gree-) GREENVILLE HOSPITAL SYSTEM.
 XX
 PI Yu X, Wagner TE;
 XX

DR WPI: 2002-154423/20.
 XX
 PT A procytotoxin comprising a cytotoxic peptide with at least one lysine
 residue which is converted into a cytotoxin in a target cell is used
 PT for treating cancer, particularly of the prostate, skin, ovary or lung
 PT
 XX
 PS Disclosure: Page 12; 33pp; English.
 XX
 CC The present invention relates to a procytotoxin, comprising a cytotoxic
 CC peptide with at least one lysine residue bound by a peptide bond to at
 CC least one amino acid via the epsilon-amino group of the lysine residue.
 CC The procytotoxin is used to treat cancer, particularly of the prostate,
 CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
 CC peptide, typically disrupt cell membrane, causing cell lysis and death
 CC (apoptosis) upon contact. The present sequence is proctoxin.
 CC
 XX
 SQ Sequence 27 AA:
 Query Match 73.6%; Score 101.5; DB 23; Length 27;
 Best Local Similarity 75.8%; Pred. No. 7.4e-12;
 Matches 25; Conservative 0; Mismatches 1; Indels 7; Gaps 2;
 OY 1 GFATLCTKXXXXVLDGIDKXXXXLIQLIEDK 33
 ||||| ||||| ||||| |||||
 DB 1 GFATLCTK----VLDGIDK---FLIQLIEDK 26
 RESULT 11
 AAO21741
 ID AAO21741 standard: Peptide: 24 AA.
 XX
 AC AAO21741:
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Procytotoxin cytolytic peptide #6.
 XX
 KW Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;
 KW ovary; prostate; breast; skin; lung; pancreas; cytolytic.
 OS
 XX
 OS Unidentified.
 XX
 PN US2002045736-A1.
 XX
 PD 18-APR-2002.
 XX
 PF 27-AUG-2001: 2001US-0938623.
 XX
 PR 09-MAY-2001: 2001US-0851422.
 XX
 PA (YUXX/) YU X.
 PA (WAGN/) WAGNER T E.
 XX
 PI Yu X, Wagner TE;
 XX
 DR WPI: 2002-507251/54.
 XX
 PT A new procytotoxin useful in the treatment of cancer of e.g. prostate,
 PT ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator
 PT via a peptide bond cleavable by a specific protease
 PT
 XX
 PS Example 1: Page 10; 21pp; English.
 XX
 CC The invention relates to a procytotoxin comprising a cytotoxic peptide
 CC bound to an inactivator via a peptide bond, where the peptide bond is
 CC susceptible to cleavage by a targeting specific protease. The
 CC procytotoxin is used to treat cancer, particularly of the prostate,
 CC ovary, breast, skin, lung or pancreas. This sequence represents a
 CC procytotoxin cytolytic peptide sequence relating to the invention.
 XX
 SQ Sequence 24 AA:

Query Match 70.7%; Score 97.5; DB 23; Length 24;
Best Local Similarity 72.7%; Pred. No. 3.7e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 9; Gaps 2;

OY 1 GFIAFLCTKXXXXVLDGIDKXXXXLIQIEDK 33
DB 1 GFIAFLCTK----VLDGIDK-----IQIEDK 24

RESULT 12
AAE18203
ID AAE18203 standard; peptide: 24 AA.
XX AAE18203;
XX
XX
XX
XX 07-MAY-2002 (first entry)
XX
XX
XX Cytolytic peptide for treating cancer.
XX
XX Procaryotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
XX prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX
XX Unidentified.
XX
XX WO200185777-A2.
XX
XX 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-US40690.
XX
XX 09-MAY-2000; 2000US-203063P.
XX
XX 16-JUN-2000; 2000US-212042P.
XX
XX
XX (GREE-) GREENVILLE HOSPITAL SYSTEM.
XX
XX Yu X, Wagner TE;
XX
XX WPI: 2002-154423/20.
XX
XX
XX A procaryotoxin comprising a cytotoxic peptide with at least one lysine
XX residue which is converted into a cytotoxin in a target cell is used
XX for treating cancer, particularly of the prostate, skin, ovary or lung
XX
XX
XX Example 1; Page 20; 33pp; English.
XX
XX The present invention relates to a procaryotoxin, comprising a cytotoxic
XX peptide with at least one lysine residue bound by a peptide bond to at
XX least one amino acid via the epsilon-amino group of the lysine residue.
XX The procaryotoxin is used to treat cancer, particularly of the prostate,
XX ovary, lung or skin. Cytolytic peptide also known as channel-forming
XX peptide, typically disrupt cell membrane, causing cell lysis and death
XX (apoptosis) upon contact. The present sequence is cytolytic peptide.
XX
XX Sequence 24 AA:
XX
XX
XX Query Match 70.7%; Score 97.5; DB 23; Length 24;
XX Best Local Similarity 72.7%; Pred. No. 3.7e-11;
XX Matches 24; Conservative 0; Mismatches 0; Indels 9; Gaps 2;

OY 1 GFIAFLCTKXXXXVLDGIDKXXXXLIQIEDK 33
DB 1 GFIAFLCTK----VLDGIDK-----IQIEDK 24

RESULT 13
AAE18204
ID AAE18204 standard; peptide: 26 AA.
XX
XX AAE18204;
XX
XX
XX 07-MAY-2002 (first entry)
XX

DE Procaryolytic peptide for treating cancer.
XX
XX Procaryotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
XX prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX
XX Unidentified.
XX
XX
XX Key Location/Qualifiers
XX Modified-site 23
XX /note= "Addition of two gamma glutamate linked side-chain
XX glutamate acid residue to the epsilon amino group of the
XX C-terminal lysine"
XX
XX WO200185777-A2.
XX
XX 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-US40690.
XX
XX 09-MAY-2000; 2000US-203063P.
XX
XX 16-JUN-2000; 2000US-212042P.
XX
XX (GREE-) GREENVILLE HOSPITAL SYSTEM.
XX
XX Yu X, Wagner TE;
XX
XX WPI: 2002-154423/20.
XX
XX
XX A procaryotoxin comprising a cytotoxic peptide with at least one lysine
XX residue which is converted into a cytotoxin in a target cell is used
XX for treating cancer, particularly of the prostate, skin, ovary or lung
XX
XX
XX Example 1; Page 20; 33pp; English.
XX
XX The present invention relates to a procaryotoxin, comprising a cytotoxic
XX peptide with at least one lysine residue bound by a peptide bond to at
XX least one amino acid via the epsilon-amino group of the lysine residue.
XX The procaryotoxin is used to treat cancer, particularly of the prostate,
XX ovary, lung or skin. Cytolytic peptide also known as channel-forming
XX peptide, typically disrupt cell membrane, causing cell lysis and death
XX (apoptosis) upon contact. The present sequence is procaryolytic peptide.
XX
XX Sequence 26 AA:
XX
XX
XX Query Match 70.7%; Score 97.5; DB 23; Length 26;
XX Best Local Similarity 72.7%; Pred. No. 4e-11;
XX Matches 24; Conservative 0; Mismatches 0; Indels 9; Gaps 2;

OY 1 GFIAFLCTKXXXXVLDGIDKXXXXLIQIEDK 33
DB 1 GFIAFLCTK----VLDGIDK-----IQIEDK 24

RESULT 14
AAE18205
ID AAE18205 standard; peptide: 26 AA.
XX
XX AAE18205;
XX
XX 07-MAY-2002 (first entry)
XX
XX Procaryotoxin #5 for treating cancer.
XX
XX Procaryotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
XX prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX
XX Unidentified.
XX
XX
XX Key Location/Qualifiers
XX Modified-site 24..26
XX /note= "Lys([epsilon]-gamma)-Glu-[alpha]-gamma)-Glu);
XX where [epsilon]-gamma) represents a peptide bond between

FT the epsilon amino group of lysine and the gamma carboxyl
FT group of adjacent glutamate, [alpha-gamma] represents a
FT peptide bond between alpha amino group of glutamate and
FT gamma carboxyl group of adjacent glutamate"

XX W0200185777-A2.
XX
XX 15-NOV-2001.
XX
XX 09-MAY-2001; 2001MO-USA0690.
XX
XX 09-MAY-2000; 20000US-203063P.
XX
XX 16-JUN-2000; 20000US-212042P.
XX
XX (GREE-) GREENVILLE HOSPITAL SYSTEM.
XX
XX Yu X, Wagner TE;
XX
XX WPI; 2002-154423/20.
XX
XX
XX A procytotoxin comprising a cytotoxic peptide with at least one lysine
XX residue which is converted into a cytotoxin in a target cell is used
XX for treating cancer, particularly of the prostate, skin, ovary or lung
XX
XX
XX Claim 8; Page 24; 33pp; English.
XX
XX The present invention relates to a procytotoxin, comprising a cytotoxic
XX peptide with at least one lysine residue bound by a peptide bond to at
XX least one amino acid via the epsilon-amino group of the lysine residue.
XX The procytotoxin is used to treat cancer, particularly of the prostate,
XX ovary, lung or skin. Cytolytic peptide also known as channel-forming
XX peptide, typically disrupt cell membrane, causing cell lysis and death
XX (apoptosis) upon contact. The present sequence is procytotoxin.
XX
XX Sequence 26 AA:
XX
XX
XX Query Match 70.7%; Score 97.5; DB 23; Length 26;
XX Best Local Similarity 72.7%; Pred. No. 4e-11;
XX Matches 24; Conservative 0; Mismatches 0; Indels 9; Gaps 2;
XX
XX Oy 1 GFATLCTKXXXXVDPGIDKXXXXLIQLEDK 33
XX 1 GFATLCTK----VDPGIDK-----IQLEDK 24
XX DB

RESULT 15
ACG41753
ID ACG41753 standard; Protein: 463 AA.
XX
XX ACG41753:
XX
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 51987.
XX
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX
XX 25-FEB-1999; 990US-0121825.
XX 05-MAR-1999; 990US-0123180.
XX 09-MAR-1999; 990US-0123548.
XX 23-MAR-1999; 990US-0125788.
XX 25-MAR-1999; 990US-0126264.

PR 29-MAR-1999; 990US-0126785.
PR 01-APR-1999; 990US-0127462.
PR 06-APR-1999; 990US-0128234.
PR 08-APR-1999; 990US-0128714.
PR 16-APR-1999; 990US-0129845.
PR 19-APR-1999; 990US-0130077.
PR 21-APR-1999; 990US-0130449.
PR 23-APR-1999; 990US-0130510.
PR 23-APR-1999; 990US-0130891.
PR 28-APR-1999; 990US-0131449.
PR 30-APR-1999; 990US-0132048.
PR 30-APR-1999; 990US-0132407.
PR 04-MAY-1999; 990US-0132484.
PR 05-MAY-1999; 990US-0132485.
PR 06-MAY-1999; 990US-0132486.
PR 06-MAY-1999; 990US-0132487.
PR 07-MAY-1999; 990US-0132863.
PR 11-MAY-1999; 990US-0134256.
PR 14-MAY-1999; 990US-0134219.
PR 14-MAY-1999; 990US-0134218.
PR 14-MAY-1999; 990US-0134219.
PR 14-MAY-1999; 990US-0134370.
PR 18-MAY-1999; 990US-0134370.
PR 19-MAY-1999; 990US-0134768.
PR 20-MAY-1999; 990US-0134941.
PR 21-MAY-1999; 990US-0135124.
PR 21-MAY-1999; 990US-0135353.
PR 24-MAY-1999; 990US-0135629.
PR 25-MAY-1999; 990US-0136021.
PR 27-MAY-1999; 990US-0136392.
PR 28-MAY-1999; 990US-0136782.
PR 01-JUN-1999; 990US-0137222.
PR 03-JUN-1999; 990US-0137528.
PR 04-JUN-1999; 990US-0137502.
PR 07-JUN-1999; 990US-0137724.
PR 08-JUN-1999; 990US-0138094.
PR 10-JUN-1999; 990US-0138540.
PR 10-JUN-1999; 990US-0138847.
PR 14-JUN-1999; 990US-0139119.
PR 16-JUN-1999; 990US-0139452.
PR 17-JUN-1999; 990US-0139453.
PR 17-JUN-1999; 990US-0139452.
PR 18-JUN-1999; 990US-0139454.
PR 18-JUN-1999; 990US-0139455.
PR 18-JUN-1999; 990US-0139456.
PR 18-JUN-1999; 990US-0139457.
PR 18-JUN-1999; 990US-0139458.
PR 18-JUN-1999; 990US-0139459.
PR 18-JUN-1999; 990US-0139460.
PR 18-JUN-1999; 990US-0139461.
PR 18-JUN-1999; 990US-0139462.
PR 18-JUN-1999; 990US-0139463.
PR 18-JUN-1999; 990US-0139750.
PR 18-JUN-1999; 990US-0139763.
PR 21-JUN-1999; 990US-0139817.
PR 22-JUN-1999; 990US-0139899.
PR 23-JUN-1999; 990US-0140033.
PR 23-JUN-1999; 990US-0140354.
PR 24-JUN-1999; 990US-0140695.
PR 28-JUN-1999; 990US-0140823.
PR 29-JUN-1999; 990US-0140991.
PR 30-JUN-1999; 990US-0141287.
PR 01-JUL-1999; 990US-0141842.
PR 01-JUL-1999; 990US-0142154.
PR 02-JUL-1999; 990US-0142055.
PR 06-JUL-1999; 990US-0142390.
PR 08-JUL-1999; 990US-0142803.
PR 09-JUL-1999; 990US-0142920.
PR 12-JUL-1999; 990US-0142977.
PR 13-JUL-1999; 990US-0143542.
PR 14-JUL-1999; 990US-0143624.
PR 15-JUL-1999; 990US-0144005.
PR 16-JUL-1999; 990US-0144086.
PR 16-JUL-1999; 990US-0144086.

PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144332.
 PR 20-JUL-1999; 99US-0144332.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 26-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 31.9%; Score 44; DB 21; Length 463;
 Best Local Similarity 47.4%; Pred. No. 11;
 Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
 Oy 2 FIATLCTKXXXXVLDPCID 20
 Db 282 FIATLCTKXXXXVLDPCID 300

Search completed: May 9, 2003, 16:15:34
 Job time : 53.2353 secs

THIS PAGE BLANK (US70)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:13:10 ; Search time 13.4216 Seconds
(Without alignments)
81.112 Million cell updates/sec

Title: US-09-851-422a-1
Perfect score: 138
Sequence: 1 GFATLCTKXXXXVLDGIDKXXXXXLIQLIEDKXXXX 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/1/1aa/Backfilltest.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	30.4	893	4	US-09-019-160-6
2	41.3	30.1	460	4	US-09-199-637A-132
3	41	29.7	893	2	US-08-706-702-3
4	41	29.7	893	3	US-08-706-706-3
5	41	29.7	893	4	US-09-019-160-2
6	41	29.7	893	4	US-09-019-160-7
7	41	29.7	893	4	US-09-019-160-8
8	41	29.7	893	4	US-09-019-160-9
9	40	29.0	144	4	US-09-134-001C-5477
10	37	26.8	40	2	US-08-807-332B-5
11	37	26.8	126	4	US-09-338-876-5
12	37	26.8	126	4	US-09-134-001C-4599
13	37	26.8	262	1	US-08-738-944-50
14	37	26.8	262	4	US-09-263-352-40
15	37	26.8	395	1	US-08-723-938-3
16	37	26.8	395	2	US-09-080-538-3
17	37	26.8	396	2	US-09-061-337-12
18	37	26.8	396	2	US-09-123-129-12
19	37	26.8	396	3	US-09-340-991-12
20	37	26.8	396	4	US-08-974-609-12
21	37	26.8	396	4	US-09-242-859A-2
22	37	26.8	396	4	US-09-242-859A-6
23	37	26.8	403	2	US-09-549-098-12
24	37	26.8	403	2	US-09-061-337-10
25	37	26.8	403	2	US-09-122-129-10
26	37	26.8	403	3	US-09-340-991-10
27	37	26.8	403	4	US-08-974-609-10

28	37	26.8	403	4	US-09-549-098-10	Sequence 10, Appl
29	37	26.8	420	4	US-09-008-271A-4	Sequence 4, Appl
30	37	26.8	420	4	US-08-974-691-8	Sequence 8, Appl
31	37	26.8	420	4	US-09-705-448-1	Sequence 1, Appl
32	36	26.1	91	4	US-09-134-001C-2969	Sequence 2969, Ap
33	36	26.1	103	2	US-08-596-319-31	Sequence 31, Appl
34	36	26.1	324	2	US-08-671-320-11	Sequence 11, Appl
35	36	26.1	324	2	US-08-868-577-11	Sequence 11, Appl
36	36	26.1	675	4	US-09-134-001C-4547	Sequence 4547, Ap
37	36	26.1	761	1	US-07-906-395-2	Sequence 2, Appl
38	36	26.1	761	1	US-08-192-632-2	Sequence 2, Appl
39	36	26.1	761	1	US-08-710-676-2	Sequence 2, Appl
40	36	26.1	761	3	US-09-099-902B-2	Sequence 2, Appl
41	36	26.1	761	5	PCT-US93-06080-2	Sequence 2, Appl
42	36	26.1	892	1	US-07-977-434-12	Sequence 12, Appl
43	36	26.1	892	1	US-08-458-819-12	Sequence 12, Appl
44	36	26.1	892	5	PCT-US91-07035-12	Sequence 12, Appl
45	35	25.4	250	4	US-09-134-001C-4399	Sequence 4399, Ap

ALIGNMENTS

RESULT 1
US-09-019-160-6
Sequence 6, Application US/09019160
Patent No. 6306588
GENERAL INFORMATION:
Applicant: Chatterjee, Deb K.
Applicant: Solus, Joseph
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
TITLE OF INVENTION: Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/019,160
FILING DATE: 06-FEB-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: (To be assigned)
FILING DATE: 06-JAN-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
US-09-019-160-6

Query Match 30.4%; Score 42; DB 4; Length 893;
Best Local Similarity 25.8%; Pred. No. 25;
Matches 8; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

Qy 3 IATLCTKXXXVLDGIDKXXXXLQLEDEK 33
Db 117 IATLAVRAARFLMRPSLTGDKMDLVNEK 147

RESULT 2
US-09-199-637A-132

Sequence 132, Application US/09199637A
Patent No. 6355411
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick
APPLICANT: Goodman, Howard M.
APPLICANT: Rahme, Laurence G.
APPLICANT: Mahajan-Miklos, Shailina
APPLICANT: Tan, Man-Wah
APPLICANT: Cao, Hui
APPLICANT: Drenkard, Eliana
APPLICANT: Tsongalis, John
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
FILE REFERENCE: 00786/361002
CURRENT APPLICATION NUMBER: US/09/199,637A
CURRENT FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: 60/066,517
PRIOR FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 132
LENGTH: 460
TYPE: PRF
ORGANISM: Pseudomonas aeruginosa
US-09-199-637A-132

Query Match 30.1%; Score 41.5; DB 4; Length 460;
Best Local Similarity 52.9%; Pred. No. 14;
Matches 9; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

Qy 17 FGIDKXXXXL-LQLEDEK 32
Db 287 FGIDKPGCALVALQVED 303

RESULT 3
US-08-706-702-3

Sequence 3, Application US/08706702
Patent No. 5948614
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
TITLE OF INVENTION: Cloned DNA polymerases from Thermotoga
TITLE OF INVENTION: maritima and Mutants Thereof
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,702
FILING DATE: 06-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/689,807
FILING DATE: 14-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,400
FILING DATE: 02-OCT-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/576,759
FILING DATE: 21-DEC-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,397
FILING DATE: 02-OCT-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/525,057
FILING DATE: 08-SEP-1995

ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942,2800006

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
TOPOLOGY: linear

US-08-706-702-3
MOLECULE TYPE: protein

Query Match 29.7%; Score 41; DB 2; Length 893;
Best Local Similarity 25.8%; Pred. No. 38;
Matches 8; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

Qy 3 IATLCTKXXXVLDGIDKXXXXLQLEDEK 33
Db 117 IATLAVRAARFLMRPSLTGDKMDLVNEK 147

RESULT 4

US-08-706-706-3

Sequence 3, Application US/08706706
Patent No. 6015668
GENERAL INFORMATION:
APPLICANT: Hughes, A. John
APPLICANT: Chatterjee, Deb K.
TITLE OF INVENTION: Cloned DNA polymerases from Thermotoga and
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/706,706
FILING DATE: 06-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/689,807
FILING DATE: 14-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,400
FILING DATE: 02-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/370,190
FILING DATE: 01-JAN-1995
PRIOR APPLICATION DATA:

Query Match	29.7%	Score 41	DB 3	Length 893:
Best Local Similarity	25.8%	Pred. No. 38:		
Matches	8:	Conservative	8:	Mismatches 15: Indels 0: Gaps 0:
Qy	3	TATLCTKXXAVLDFGIDKXXXXLQIILEDK	33	
		:::::		
Db	117	TATLAVRAARPLMRFSLTGDKDMQLVNEK	147	

```

1 REFERENCE/DOCKET NUMBER: 09442.4250002
2
3 TELECOMMUNICATION INFORMATION:
4
5 TELEPHONE: 202-371-2600
6
7 TELEFAX: 202-371-2540
8
9 INFORMATION FOR SEQ ID NO: 2:
10
11 SEQUENCE CHARACTERISTICS:
12
13 LENGTH: 893 amino acids
14
15 TYPE: amino acid
16
17 STRANDEDNESS: not relevant
18
19 TOPOLOGY: not relevant
20
21 MOLECULE TYPE: protein
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1
```

Query Match	29.7%	Score 41	DB	4	Length 893
Best Local Similarity	25.8%	Pred. No.	38		
Matches	8	Conservative	8	Mismatches	15
				Indels	0
				Gaps	0
QY	3	IATLCTKXXXVLPFGIDKXXXXLIIQLIDK	33		
		:::::	:::::		
Db	117	IATLAVRAARFLMRESLTGDKDMQLVNEK	147		

RESULT 6
 US-09-019-160-7
 Sequence 7, Application US/09019160
 Patent No. 6306588
 GENERAL INFORMATION:
 APPLICANT: Chatterjee, Deb K.
 APPLICANT: Solus, Joseph
 APPLICANT: Yang, Shuwei
 TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
 TITLE OF INVENTION: Nucleic Acid Fragments and Uses Thereof
 NUMBER OF SEQUENCES: 93
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C
 STREET: 1100 New York Ave., N.W., Suite 600
 City: Washington
 STATE: DC
 COUNTRY: USA
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/019,160
 FILING DATE: 06-FEB-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: (To be assigned)
 FILING DATE: 06-JAN-1998
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/037,393
 FILING DATE: 07-FEB-1997
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 0942,4250002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-371-2600
 TELEFAX: 202-371-2540
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 893 amino acids
 TYPE: amino acid
 STRANDEDNESS: not relevant
 TOPOLOGY: not relevant
 MOLECULE TYPE: protein
 US-09-019-160-7

Query Match 29.7% Score 41; DB 4; Length 893;
Best Local Similarity 25.8%; Pred. No. 38;
Matches 8; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 3 IATLCTKXXXXVLDGIDKXXXXLIIOLIEDK 33

DB 117 IATLAVRAARFLMRFSLTIGDKMDQLVNEK 147

RESULT 7
US-09-019-160-8
Sequence 8, Application US/09019160

Patent No. 6306588
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
APPLICANT: Solus, Joseph
APPLICANT: Yang, Shuwei
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
TITLE OF INVENTION: Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNF. KESSLER, GOLDSTEIN & FOX, P.L.L.C
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/019,160
FILING DATE: 06-FEB-1998
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: (To be assigned)
FILING DATE: 06-JAN-1998
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-019-160-8

Query Match 29.7% Score 41; DB 4; Length 893;
Best Local Similarity 25.8%; Pred. No. 38;
Matches 8; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 3 IATLCTKXXXXVLDGIDKXXXXLIIOLIEDK 33

DB 117 IATLAVRAARFLMRFSLTIGDKMDQLVNEK 147

RESULT 8
US-09-019-160-9
Sequence 9, Application US/09019160
Patent No. 6306588

GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
APPLICANT: Solus, Joseph
APPLICANT: Yang, Shuwei
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
TITLE OF INVENTION: Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNF. KESSLER, GOLDSTEIN & FOX, P.L.L.C
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/019,160
FILING DATE: 06-FEB-1998
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: (To be assigned)
FILING DATE: 06-JAN-1998
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-09-019-160-9

Query Match 29.7% Score 41; DB 4; Length 893;
Best Local Similarity 25.8%; Pred. No. 38;
Matches 8; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 3 IATLCTKXXXXVLDGIDKXXXXLIIOLIEDK 33

DB 117 IATLAVRAARFLMRFSLTIGDKMDQLVNEK 147

RESULT 9
US-09-134-001C-5477
Sequence 5477, Application US/09134001C
Patent No. 6380370

GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIORITY FILING DATE: 1998-08-13
PRIORITY APPLICATION NUMBER: US 60/064,964
PRIORITY FILING DATE: 1997-11-08
PRIORITY APPLICATION NUMBER: US 60/055,779
PRIORITY FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5477


```
; LENGTH: 144
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5477

Query Match          29.0%; Score 40; DB 4; Length 144;
Best Local Similarity 41.2%; Pred. No. 6.9;
Matches 7; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 5 TLCTKXXXXVLDGIDK 21
   1 111 :1:1
Db 96 TYCTKIIYQAYKYGVK 112

RESULT 10
US-08-807-332B-5
; Sequence 5, Application US/08807332B
; Patent No. 5959074
; GENERAL INFORMATION:
; APPLICANT: Dreyfus, David H.
; APPLICANT: Gelfand, Erwin W.
; TITLE OF INVENTION: PRODUCTS AND PROCESSES FOR REGULATION OF
; TITLE OF INVENTION: GENE RECOMBINATION
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/807,332B
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kovarik, Joseph E.
; REGISTRATION NUMBER: 33,005
; REFERENCE/DOCKET NUMBER: 2879-39
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-807-332B-5

Query Match          26.8%; Score 37; DB 2; Length 40;
Best Local Similarity 36.8%; Pred. No. 5.5;
Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

OY 1 GFIATLCTKXXXXVLDGFI 19
   1 1111 :1:1
Db 17 GHYOTLCTNFRSLAIDKGV 35

RESULT 11
US-09-338-876-5
; Sequence 5, Application US/09338876
; Patent No. 6187584
; GENERAL INFORMATION:
; APPLICANT: Dreyfus, David H.
; APPLICANT: Gelfand, Erwin W.
; TITLE OF INVENTION: PRODUCTS AND PROCESSES FOR REGULATION OF
```

```
; TITLE OF INVENTION: GENE RECOMBINATION
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/338,876
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/807,332
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kovarik, Joseph E.
; REGISTRATION NUMBER: 33,005
; REFERENCE/DOCKET NUMBER: 2879-39
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-338-876-5

Query Match          26.8%; Score 37; DB 4; Length 40;
Best Local Similarity 36.8%; Pred. No. 5.5;
Matches 7; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

OY 1 GFIATLCTKXXXXVLDGFI 19
   1 1111 :1:1
Db 17 GHYOTLCTNFRSLAIDKGV 35

RESULT 12
US-09-134-001C-4599
; Sequence 4599, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4599
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4599

Query Match          26.8%; Score 37; DB 4; Length 126;
Best Local Similarity 31.0%; Pred. No. 21;
Matches 9; Conservative 4; Mismatches 16; Indels 0; Gaps 0;

OY 1 GFIATLCTKXXXXVLDGIDKXXXXLIQL 29
```

DB 43 GIATMPAVYLAIALALIDHKCOOLIM 71

RESULT 13

US-08-738-944-50
Sequence 50, Application US/08738944

Patent No. 5783431

GENERAL INFORMATION:

APPLICANT: Peterson, et al.

TITLE OF INVENTION: METHODS FOR GENERATING AND

NUMBER OF SEQUENCES: 51 SCREENING NOVEL METABOLIC PATHWAYS

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036/2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/738,944

FILING DATE: 24-OCT-1996

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/639,255

FILING DATE: 24-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 8757-007

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 262 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: CXC-AMN20

LOCATION: 1...262

OTHER INFORMATION:

US-08-738-944-50

Query Match 26.8%; Score 37; DB 1; Length 262;

Best Local Similarity 36.7%; Pred. No. 50;

Matches 11; Conservative 4; Mismatches 13; Indels 2; Gaps 1;

QY 4 ATLCRXXXXXVDFGIDKXXXXLIQLEDK 33

DB 226 ATLRTKFGKAYODF--DTISISYLEMMIDK 253

RESULT 14

US-09-263-352-40

Sequence 40, Application US/09263352

Patent No. 6242211

GENERAL INFORMATION:

APPLICANT: Peterson, T.

APPLICANT: Brian, P.

TITLE OF INVENTION: METHODS FOR GENERATING AND SCREENING NOVEL METABOLIC

NUMBER OF SEQUENCES: 51

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036/2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/738,944

FILING DATE: 24-OCT-1996

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/639,255

FILING DATE: 24-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 8757-007

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 50:

SEQUENCE CHARACTERISTICS:

LENGTH: 262 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: CXC-AMN20

LOCATION: 1...262

OTHER INFORMATION:

US-08-738-944-50

Query Match 26.8%; Score 37; DB 1; Length 262;

Best Local Similarity 36.7%; Pred. No. 50;

Matches 11; Conservative 4; Mismatches 13; Indels 2; Gaps 1;

QY 4 ATLCRXXXXXVDFGIDKXXXXLIQLEDK 33

DB 226 ATLRTKFGKAYODF--DTISISYLEMMIDK 253

RESULT 15

US-08-723-938-3

Sequence 3, Application US/08723938

Patent No. 576759

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Coleman, Roger

TITLE OF INVENTION: TWO NOVEL HUMAN CATHESPIN PROTEINS

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: U.S.

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/723,938

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0125 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 395 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

IMMEDIATE SOURCE:

LIBRARY: LUNGNOT02

CLONE: 312099

US-08-723-938-3

Query Match 26.8%; Score 37; DB 1; Length 395;

Best Local Similarity 36.7%; Pred. No. 50;

Matches 11; Conservative 4; Mismatches 13; Indels 2; Gaps 1;

QY 4 ATLCRXXXXXVDFGIDKXXXXLIQLEDK 33

DB 226 ATLRTKFGKAYODF--DTISISYLEMMIDK 253

Best Local Similarity 50.0%; Pred. No. 80;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 5 TLCTKXXXVLDG 18
||| |
:| |
Db 272 TLCAKGCNAILDG 285

Search completed: May 9, 2003, 16:19:26
Job time : 15.4216 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: May 9, 2003, 16:15:46 ; Search time 15.598 Seconds
(without alignments)
218.294 Million cell updates/sec

Title: US-09-851-422a-1
Perfect score: 138
Sequence: 1 GFATLCTKXXVYVLDGIDKXXXXLQLEDKXXXX 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_MA:*

- 1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	31.2	431	9	US-09-986-480-189
2	42	30.4	893	9	US-09-891-332A-6
3	41.5	30.1	460	9	US-09-975-719-132
4	41	29.7	893	9	US-09-891-332A-7
5	41	29.7	893	9	US-09-891-332A-8
6	41	29.7	893	9	US-09-891-332A-9
7	41	29.7	893	9	US-09-891-332A-9
8	41	29.7	893	9	US-09-229-173-3
9	41	29.7	893	10	US-09-741-664-1
10	41	29.7	893	10	US-09-741-664-2
11	40	29.0	453	10	US-09-800-729-83
12	39	28.3	718	10	US-09-928-175-21
13	39	28.3	737	9	US-10-228-735-3
14	39	28.3	737	10	US-09-928-175-20
15	37	26.8	48	10	US-09-864-761-48379
16	37	26.8	172	9	US-09-860-670-112
17	37	26.8	212	10	US-09-915-582-87
18	37	26.8	285	10	US-09-925-302-670
19	37	26.8	317	10	US-09-815-242-13183

20	37	26.8	390	9	US-09-969-384-15	Sequence 15, Appl
21	37	26.8	395	9	US-10-094-080-3	Sequence 3, Appl
22	37	26.8	433	9	US-09-964-899-11	Sequence 11, Appl
23	36	26.1	58	9	US-10-091-504-967	Sequence 967, App
24	36	26.1	58	10	US-09-764-869-967	Sequence 967, App
25	36	26.1	155	1	US-08-781-966A-5255	Sequence 5255, Ap
26	36	26.1	213	10	US-09-844-468-4	Sequence 4, Appl
27	36	26.1	436	10	US-09-796-766-21	Sequence 21, Appl
28	36	26.1	618	9	US-09-738-626-3975	Sequence 3975, Ap
29	36	26.1	1192	9	US-10-170-102-2	Sequence 2, Appl
30	35	25.4	116	10	US-09-864-761-34645	Sequence 34645, A
31	35	25.4	131	10	US-09-881-752A-326	Sequence 326, App
32	35	25.4	283	9	US-09-738-626-3533	Sequence 3533, Ap
33	35	25.4	381	9	US-09-738-626-5896	Sequence 5896, Ap
34	35	25.4	724	9	US-09-881-579-4	Sequence 4, Appl
35	35	25.4	856	10	US-09-815-242-11489	Sequence 11489, A
36	35	25.4	885	10	US-09-828-423-5	Sequence 5, Appl
37	34.5	25.0	315	10	US-09-815-242-13514	Sequence 13514, A
38	34.5	25.0	372	10	US-09-815-242-5277	Sequence 5277, Ap
39	34.5	25.0	384	10	US-09-815-242-12605	Sequence 12605, A
40	34.5	25.0	489	9	US-09-834-721-2	Sequence 2, Appl
41	34.5	25.0	489	9	US-09-834-721-4	Sequence 4, Appl
42	34.5	25.0	489	9	US-09-783-388-2	Sequence 2, Appl
43	34.5	25.0	489	9	US-09-783-388-4	Sequence 4, Appl
44	34.5	25.0	489	9	US-09-951-535-2	Sequence 2, Appl
45	34.5	25.0	489	9	US-09-951-535-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-986-480-189
Sequence 189, Application US/09986480
Publication No. US20030027899A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 143 Human secreted proteins
FILE REFERENCE: PSS00p1
CURRENT APPLICATION NUMBER: US/09/986, 480
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: PCT/US00/12788
PRIOR FILING DATE: 2000-05-11
PRIOR APPLICATION NUMBER: US 60/134, 068
PRIOR FILING DATE: 1999-05-13
NUMBER OF SEQ ID NOS: 456
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 189
LENGTH: 431
TYPE: PRT
ORGANISM: Homo sapiens
US-09-986-480-189

Query Match 31.2% Score 43. DB 9; Length 431;
Best Local Similarity 45.5% Pred. No. 11;
Matches 15; Conservative 1; Mismatches 15; Indels 2; Gaps 1;

QY 3 IATLCT--KXXXXVLDGIDKXXXXLQLEDK 33
DB 13 IATLCTLKEPYAVRYRGEXKDNALIAOLIDK 45

RESULT 2
US-09-891-332A-6
Sequence 6, Application US/09891332A
Patent No. US2002016646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Soltus, Joseph
Yang, Shuwei
TITLE OF INVENTION: Polymers for Analyzing or Typing Polymorphic
Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93

```

CORRESPONDENCE ADDRESS:
ADDRESS: STERNF, KESSLER, GOLDSTEIN & FOX, P.L.L.C
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/891,332A
FILING DATE: 27-Jun-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,160
FILING DATE: <unknown>
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942,4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-891-332A-6.

Query Match      30.1% Score 41.5 DB 9 Length 893
Best Local Similarity 25.8% Pred. No. 37
Matches 8: Conservative 8 Mismatches 15 Indels 0 Gaps 0
OY 3 IATLCTKXXXXVLDPGIDKXXXXXLIQLIEDK 33
DB 117 IATLAVRAARFLMRSLITGAKDMQLVNEK 147

RESULT 3
US-09-975-719-132
Sequence 132, Application US/09975719
Publication No. US20030022349A1
GENERAL INFORMATION:
APPLICANT: Ausubel, Frederick M.
TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
FILE REFERENCE: 00786/361003
CURRENT FILING DATE: US/09/975,719
CURRENT FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 09/199,637
PRIOR FILING DATE: 1998-11-25
PRIOR APPLICATION NUMBER: US 60/066,517
PRIOR FILING DATE: 1997-11-25
NUMBER OF SEQ ID NOS: 437
SOFTWARE: PatSeq for Windows Version 4.0
SEQ ID NO 132
LENGTH: 460
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-975-719-132

Query Match      30.1% Score 41.5 DB 9 Length 460
Best Local Similarity 25.9% Pred. No. 22

```

```

Matches 9: Conservative 3: Mismatches 4: Indels 1: Gaps
QY 17 FGIIDKXXXXLI-QLIID 32
||:|:| |:|:|:|
Db 287 FGIDKPGSALVAOLVED 303

RESULT 4
US-09-891-332A-2
; Sequence 2, Application US/09891332A
; Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Solus, Joseph
Yang, Shuwei
TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic
Nucleic Acid Fragments and Uses Thereof
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNER, KESSLER, GOLDSTEIN & FOX, P.L.L.C
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/891,332A
FILING DATE: 27-Jun-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,160
FILING DATE: <Unknown>
APPLICATION NUMBER: US 60/037,393
FILING DATE: 07-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.4250002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-891-332A-2

Query Match 29.7%: Score 41: DB 9: Length 893:
Best Local Similarity 25.8%: Pred. No. 55:
Matches 8: Conservative 8: Mismatches 15: Indels 0: Gaps 0:
QY 3 IATLCTKXXXXXVADFGIDKXXXXXXLIQLIEDK 33
|||:|:|:|:|:|:|:|:|
Db 117 IATLAVRAARFLMRFSLTITGDKDMQIVNEK 147

RESULT 5
US-09-891-332A-7
; Sequence 7, Application US/09891332A
; Patent No. US20020168646A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
Solus, Joseph
Yang, Shuwei

```

TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic Nucleic Acid Fragments and Uses Thereof

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESSES:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.

STREET: 1100 New York Ave., N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/891.332A

FILING DATE: 27-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/019,160

FILING DATE: <Unknown>

APPLICATION NUMBER: US 60/037,393

FILING DATE: 07-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Esmond, Robert W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0942.4250002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 893 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-891-332A-7

Query Match 29.7% Score 41; DB 9; Length 893;

Best Local Similarity 25.8% Pred. No. 55;

Matches 8; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

OY 3 IATLCTKXXVYLDGIDKXXXXXLIQLEDK 33

DB 117 IATLAVRAARFLMRSLITGDKMDQLVNER 147

RESULT 6

US-09-891-332A-8

Sequence 8, Application US/09891332A

Patent No. US20020168646A1

GENERAL INFORMATION:

APPLICANT: Chatterjee, Deb K.

Solus, Joseph

Yang, Shuwel

TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic Nucleic Acid Fragments and Uses Thereof

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESSES:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.

STREET: 1100 New York Ave., N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/891.332A

FILING DATE: 27-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/019,160

FILING DATE: <Unknown>

APPLICATION NUMBER: US 60/037,393

FILING DATE: 07-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Esmond, Robert W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0942.4250002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 893 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-09-891-332A-8

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/891.332A

FILING DATE: 27-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/019,160

FILING DATE: <Unknown>

APPLICATION NUMBER: US 60/037,393

FILING DATE: 07-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Esmond, Robert W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0942.4250002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 893 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 8:

US-09-891-332A-8

Query Match 29.7% Score 41; DB 9; Length 893;

Best Local Similarity 25.8% Pred. No. 55;

Matches 8; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

OY 3 IATLCTKXXVYLDGIDKXXXXXLIQLEDK 33

DB 117 IATLAVRAARFLMRSLITGDKMDQLVNER 147

RESULT 7

US-09-891-332A-9

Sequence 9, Application US/09891332A

Patent No. US20020168646A1

GENERAL INFORMATION:

APPLICANT: Chatterjee, Deb K.

Solus, Joseph

Yang, Shuwel

TITLE OF INVENTION: Polymerases for Analyzing or Typing Polymorphic Nucleic Acid Fragments and Uses Thereof

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESSES:

ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.

STREET: 1100 New York Ave., N.W., Suite 600

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-3934

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/891.332A

FILING DATE: 27-Jun-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/019,160

FILING DATE: <Unknown>

APPLICATION NUMBER: US 60/037,393

FILING DATE: 07-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Esmond, Robert W.

REGISTRATION NUMBER: 32,893

REFERENCE/DOCKET NUMBER: 0942.4250002

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-371-2600

TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-891-332A-9

Query Match 29.7%; Score 41; DB 9; Length 893;
Best Local Similarity 25.8%; Pred. No. 55;
Matches 8; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

OY 3 IATLCTKXXXVLDGIDKXXXLIQIIEDK 33
DB 117 IATLAVRAARFLMRPSLTGDKMDLQVNEK 147

RESULT 8
US-09-229-173-3
Sequence 3, Application US/09229173
Publication No. US20030027296A1
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
TITLE OF INVENTION: Cloned DNA Polymerases from Thermotoga
TITLE OF INVENTION: maritima and Mutants Thereof
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/229,173
FILING DATE: 13-JAN-1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/706,702
FILING DATE: 06-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/689,807
FILING DATE: 14-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,400
FILING DATE: 02-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/576,759
FILING DATE: 21-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,397
FILING DATE: 02-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/525,057
FILING DATE: 08-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Miltonig, Robert C.
REGISTRATION NUMBER: 34,395
REFERENCE/DOCKET NUMBER: 0942.2800008
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-229-173-3

Query Match 29.7%; Score 41; DB 9; Length 893;
Best Local Similarity 25.8%; Pred. No. 55;
Matches 8; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

OY 3 IATLCTKXXXVLDGIDKXXXLIQIIEDK 33
DB 117 IATLAVRAARFLMRPSLTGDKMDLQVNEK 147

RESULT 9
US-09-741-664-1
Sequence 1, Application US/09741664
Patent No. US20010041334A1
GENERAL INFORMATION:
APPLICANT: Rashchian, Ayoub
TITLE OF INVENTION: Stable Compositions for Nucleic Acid
TITLE OF INVENTION: Sequencing and Amplification
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/741,664
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/049,021
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.3910000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-741-664-1

Query Match 29.7%; Score 41; DB 10; Length 893;
Best Local Similarity 25.8%; Pred. No. 55;
Matches 8; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

OY 3 IATLCTKXXXVLDGIDKXXXLIQIIEDK 33
DB 117 IATLAVRAARFLMRPSLTGDKMDLQVNEK 147

RESULT 10
US-09-741-664-2
Sequence 2, Application US/09741664
Patent No. US20010041334A1
GENERAL INFORMATION:


```

APPLICANT: Rashchian, Ayoub
APPLICANT: Solus, Joseph
TITLE OF INVENTION: Stable Compositions for Nucleic Acid
TITLE OF INVENTION: Sequencing and Amplification
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESS: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 New York Avenue, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Ver. 1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/741,664
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/049,021
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942,3910000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2540
TELEFAX: 202-371-2600
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 893 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-741-664-2

Query Match          29.7%: Score 41; DB 10; Length 893;
Best Local Similarity 25.8%: Pred. No. 55;
Matches 8; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 3 IATLCTKXXXXVDFGIDKXXXXLQIOLIEDK 33
DB 117 IATLAVRARARFLMRSLITGDKMDQLVNEK 147

RESULT 11
US-09-800-729-83
Sequence 83, Application US/09800729
Patent No. US20020068319A1
GENERAL INFORMATION:
APPLICANT: NI et al.
TITLE OF INVENTION: 32 Human secreted proteins
FILE REFERENCE: P2044P1
CURRENT APPLICATION NUMBER: US/09/800,729
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: PCT/US00/26013
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 60/155,709
PRIOR FILING DATE: 1999-09-24
NUMBER OF SEQ ID NOS: 217
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 83
LENGTH: 453
TYPE: PRT
ORGANISM: Homo sapiens
US-09-800-729-83

Query Match          29.0%: Score 40; DB 10; Length 453;
Best Local Similarity 36.0%: Pred. No. 40;

```

```

Matches 9; Conservative 3; Mismatches 13; Indels 0; Gaps 0;

QY 5 TLCTKXXXXVDFGIDKXXXXLIOL 29
DB 241 TQCSKDKLINDLCIDRSSYTLVDL 265

RESULT 12
US-09-928-175-21
Sequence 21, Application US/09928175
Patent No. US20020123618A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher J.
APPLICANT: Gong, Jianhua
APPLICANT: Daugherty, Betsy
APPLICANT: Rogers, No. US20020123618A1ma
TITLE OF INVENTION: Leucine-rich G Protein Coupled Receptor-8 Molecules and
FILE REFERENCE: 00-1229
CURRENT APPLICATION NUMBER: US/09/928,175
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/224,455
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 21
LENGTH: 718
TYPE: PRT
ORGANISM: Mus musculus
US-09-928-175-21

Query Match          28.3%: Score 39; DB 10; Length 718;
Best Local Similarity 47.1%: Pred. No. 1e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 GFATLCTKXXXXVDF 17
DB 463 GFATLSTEVSVLLTF 479

RESULT 13
US-10-229-735-3
Sequence 3, Application US/10229735
Publication No. US20030082650A1
GENERAL INFORMATION:
APPLICANT: Baylor College of Medicine
APPLICANT: Agoulnik, Alexander I.
TITLE OF INVENTION: The GREAT Gene and Protein
FILE REFERENCE: 7572/73263
CURRENT APPLICATION NUMBER: US/10/229,735
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: 60/315,696
PRIOR FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: 60/351,432
PRIOR FILING DATE: 2002-01-28
NUMBER OF SEQ ID NOS: 60
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 737
TYPE: PRT
ORGANISM: Mus musculus
US-10-229-735-3

Query Match          28.3%: Score 39; DB 9; Length 737;
Best Local Similarity 47.1%: Pred. No. 1e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 GFATLCTKXXXXVDF 17
DB 482 GFATLSTEVSVLLTF 498

RESULT 14

```

```

US-09-928-175-20
; Sequence 20, Application US/09928175
; Patent No. US20020123618A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher J.
; APPLICANT: Gong, Jianhua
; APPLICANT: Daugherty, Betsy
; APPLICANT: Rogers, No. US20020123618A1a
; TITLE OF INVENTION: Leucine-Rich G Protein Coupled Receptor-8 Molecules and
; FILE REFERENCE: Uses Thereof
; FILE REFERENCE: 00-1229
; CURRENT APPLICATION NUMBER: US/09/928,175
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/224,455
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 20
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-928-175-20

```

```

Query Match      28.3% Score 39; DB 10; Length 737;
Best Local Similarity 47.1%; Pred. No. 1e+02;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 1 CFIATLCTKXXXXVDF 17
DB 482 GFLATLSTFVSVLLTF 498

```

```

RESULT 15
US-09-864-761-48379
; Sequence 48379, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 48379
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC008655.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 59
; OTHER INFORMATION: SWISSPROT HIT: O96009, EVALU 7.00e-22
; OTHER INFORMATION: EST_HUMAN HIT: A1924523.1, EVALU 9.00e-21
US-09-864-761-48379

```

```

Query Match      26.8% Score 37; DB 10; Length 48;
Best Local Similarity 50.0%; Pred. No. 12;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY 5 TLCTKXXXXVDFG 18
DB 8 TLCARCAAILDTG 21

```

Search completed: May 9, 2003, 16:20:16
Job time: 17.598 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:12:15 ; Search time 23.2157 seconds
(without alignments)
153.214 Million cell updates/sec

Title: US-09-851-422A-1
Perfect score: 138
Sequence: 1 GFIAFLCTKXXXXVLDGFDKXXXXLIQLIEDKXXXX 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	102	73.9	98	1 S25283	amoeba A precursor
2	63	45.7	96	1 S49144	amoeba B precursor
3	54	39.1	101	1 S49145	amoeba C precursor
4	44	31.9	215	1 A81410	hypothetical prote
5	44	31.9	607	1 HNM2PD	hemagglutinin - ph
6	44	31.9	607	1 J01535	hemagglutinin - ph
7	43	31.2	246	2 T32510	hypothetical prote
8	43	31.2	389	2 T18111	probable UDPglucos
9	43	31.2	594	2 A53455	synthaxin-binding p
10	42	30.4	444	2 H83624	hypothetical prote
11	42	30.4	498	2 C64476	serine proteinase
12	41.5	30.1	474	2 F83550	hypothetical prote
13	41	29.7	222	2 H80208	hypothetical prote
14	41	29.7	666	2 D42510	Oil protein - vacc
15	41	29.7	666	2 T28491	hypothetical prote
16	41	29.7	666	2 C72157	SIL protein - vari
17	41	29.7	666	2 F36842	RNA-directed DNA p
18	41	29.7	1094	2 T00814	hypothetical prote
19	40.5	29.3	674	2 T23235	hypothetical prote
20	40	29.0	217	2 G70378	hypothetical prote
21	40	29.0	487	2 B95059	type 1 site-specif
22	40	28.0	487	2 A97928	holliday junction
23	40	28.0	1007	2 T01437	hypothetical prote
24	39.5	28.6	1603	2 A82093	SWF8a protein - sw
25	39	28.3	185	1 WZVSW	probable periplasm
26	39	28.3	223	2 F81120	hypothetical prote
27	39	28.3	283	2 S01072	hypothetical prote
28	39	28.3	290	2 G36789	hypothetical prote
29	39	28.3	341	2 A82129	fatty acid/phospho

30	39	28.3	396	1 G64313	protein-export mem
31	39	28.3	1612	2 S59969	DNA topoisomerase
32	39	28.3	1626	2 A39242	DNA topoisomerase
33	38.5	27.9	278	2 S17695	chlorophyll a/b-b1
34	38.5	27.9	586	2 T51211	hypothetical prote
35	38	27.5	131	2 AD2428	hypothetical prote
36	38	27.5	153	2 B64640	acyl carrier prote
37	38	27.5	207	2 D64040	hypothetical prote
38	38	27.5	223	2 S43338	protease inhibit
39	38	27.5	234	2 T39032	hypothetical ankyr
40	38	27.5	333	2 H96814	hypothetical prote
41	38	27.5	334	2 E90037	conserved hypothet
42	38	27.5	356	2 B90387	hypothetical prote
43	38	27.5	495	2 T46700	hypothetical prote
44	38	27.5	574	2 F64414	hypothetical prote
45	38	27.5	1272	2 C64513	hypothetical prote

ALIGNMENTS

RESULT 1

S25283
amoeba A precursor - Entamoeba histolytica
N:Alternate names: pore-forming protein
C:Species: Entamoeba histolytica
C>Date: 22-Nov-1993 #sequence_revision 09-Mar-1996 #text_change 18-Jun-1999
C:Accession: S25283; A41279
R:Leipe, M.; Tannich, E.; Nickel, R.; van der Goot, G.; Patus, F.; Horstmann, R.D.;
EMBO J. 11, 3501-3506, 1992
A>Title: Primary and secondary structure of the pore-forming peptide of pathogenic En
A:Reference number: S25283; MUID:93010939; PMID:1396552
A:Accession: S25283

A:Molecule type: DNA
A:Residues: 1-98 <LE1>
A:Cross-references: EMBL:M83945; NID:g158968; PIDN:AA29111.1; PID:g158969
A>Note: parts of this sequence, including the amino end of the mature protein, were d
R:Leipe, M.; Ebel, S.; Schoenberger, O.L.; Horstmann, R.D.; Mueller-Eberhard, H.J.
Proc. Natl. Acad. Sci. U.S.A. 88, 7659-7663, 1991
A>Title: Pore-forming peptide of pathogenic Entamoeba histolytica.
A:Reference number: A41279; MUID:91352048; PMID:1881907
A:Accession: A41279
A:Molecule type: protein
A:Residues: 22-25,'X',27-28,'X',30-41,'X',43-46 <LEW>
C:Superfamily: amoeba A precursor; saposin repeat homology
C:Keywords: antibacterial
F:1-21/Domain: signal sequence #status predicted <SIG>
F:18-98/Domain: saposin repeat homology <SAP>
F:22-98/Product: amoeba A #status experimental <MAT>
F:26-98,29-92,56-67/Disulfide bonds: #status predicted

Query Match 73.9% Score 102; DB 1; Length 98;
Best Local Similarity 75.8% Pred. No. 4.1e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 8; Gaps 2;

OY 1 GFIAFLCTKXXXXVLDGFDKXXXXLIQLIEDK 33
DB 61 GFIAFLCTK-----VLDGFDK-----LIQLIEDK 85

RESULT 2

S49144
amoeba B precursor - Entamoeba histolytica
C:Species: Entamoeba histolytica
C>Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 18-Jun-1999
C:Accession: S61440; S61452; S49144
R:Leipe, M.; Andree, J.; Nickel, R.; Tannich, E.; Mueller-Eberhard, H.J.
Mol. Microbiol. 14, 895-904, 1994
A>Title: Amoebapores, a family of membranolytic peptides from cytoplasmic granules of
A:Reference number: S61439; MUID:95231296; PMID:7715451
A:Accession: S61440
A:Molecule type: DNA
A:Residues: 1-96 <LE2>

A:Cross-references: EMBL:X76904; NID:9509761; PIDN:CA454226.1; PID:9509762
A:Accession: S61452
A:Molecule type: protein
A:Residues: 20-63 <LEI>
C:Superfamily: amoebapore: saposin repeat homology
C:Keywords: antibacterial
F:1-19/Domain: signal sequence #status predicted <STIC>
F:16-96/Domain: saposin repeat homology <SAP>
F:20-96/Product: amoebapore B #status predicted <MAT>
F:24-96:27-90:54-65/Disulfide bonds: #status predicted

Query Match 45.7% Score 63; DB 1; Length 96;
Best Local Similarity 43.8% Pred. No. 0.00049;
Matches 14; Conservative 7; Mismatches 3; Indels 8; Gaps 2;

OY 1 GFATLCTKXXXXVLDGIDKXXXXLIQLI 32
DB 59 GFLGLTCK-----ILSFVDE----LVKLIEN 82

RESULT 3

amoebapore C precursor - Entamoeba histolytica
S49145
C:Species: Entamoeba histolytica
C:Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 18-Jun-1999
C:Accession: S61439; S61453; S49145
R:Leippe, M.; Andrae, J.; Mickel, R.; Tannich, E.; Mueller-Eberhard, H.J.
Mol. Microbiol. 14, 895-904, 1994
A:Title: Amoebapore, a family of membranolytic peptides from cytoplasmic granules of E.
A:Reference number: S61439; MUID:95231296; PMID:7715451
A:Accession: S61439
A:Molecule type: DNA
A:Residues: 1-101 <LE2>
A:Cross-references: EMBL:X76903; NID:9509763; PIDN:CA454225.1; PID:9509764
A:Accession: S61453
A:Molecule type: protein
A:Residues: 25-69 <LEI>
C:Superfamily: amoebapore: saposin repeat homology
C:Keywords: antibacterial
F:1-24/Domain: signal sequence #status predicted <STIC>
F:21-101/Domain: saposin repeat homology <SAP>
F:25-101/Product: amoebapore A #status experimental <MAT>
F:29-101:32-95:59-70/Disulfide bonds: #status predicted

Query Match 39.1% Score 54; DB 1; Length 101;
Best Local Similarity 35.7% Pred. No. 0.022;
Matches 10; Conservative 7; Mismatches 7; Indels 4; Gaps 1;

OY 1 GFATLCTKXXXXVLDGIDKXXXXLIQ 28
DB 64 GLVETLCTK-----IVSYGIDKLEIKILE 87

RESULT 4

AB1410
hypothetical protein Cj0618 [imported] - Campylobacter jejuni (strain NCTC 11168)
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C:Accession: AB1410
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kelley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajadaram, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrett
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyf
A:Reference number: AB1250; MUID:20150912; PMID:10688204
A:Accession: AB1410

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-215 <PAR>
A:Cross-references: GB:AL139075; GB:AL111168; NID:96967817; PIDN:CA875254.1; PID:9696808
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj0618

Query Match 31.9% Score 44; DB 2; Length 215;
Best Local Similarity 40.0% Pred. No. 3.2;
Matches 8; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 1 GFATLCTKXXXXVLDGID 20
DB 8 GTLANLCAMIAWPSIDFIID 27

RESULT 5

HMNZPD
hemagglutinin - phocine distemper virus
C:Species: phocine distemper virus
C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 16-Jul-1999
C:Accession: J01369
R:Koeneva, J.; Blixenkron-Moeller, M.; Sharma, B.; Oerell, C.; North, E.
J. Gen. Virol. 72, 2959-2966, 1991
A:Title: The nucleotide sequence and deduced amino acid composition of the haemagglut
A:Reference number: J01368; MUID:92113538; PMID:1765768
A:Accession: J01369
A:Molecule type: genomic RNA
A:Residues: 1-607 <KOY>
A:Cross-references: GB:Z36979; NID:9536790; PIDN:CA485428.1; PID:9536791
C:Genetics:
A:Gene: H
C:Superfamily: measles virus hemagglutinin
C:Keywords: glycoprotein; hemagglutinin; transmembrane protein
F:39-55/Domain: transmembrane #status predicted <TMN>
F:19,149,276,391,422,456,587/Binding site: carbohydrate (asn) (covalent) #status pred

Query Match 31.9% Score 44; DB 1; Length 607;
Best Local Similarity 28.6% Pred. No. 9.4;
Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

OY 3 IATLCTKXXXXVLDGIDKXXXXLIQLI 30
DB 292 LASLCTKESTILNLGDESONSVLVI 319

RESULT 6

J01535
hemagglutinin - phocine distemper virus (strain Ulster/88)
N:Alternate names: attachment protein
C:Species: phocine distemper virus
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Sep-1994
C:Accession: J01535
R:Curran, M.D.; O'Loan, D.; Kennedy, S.; Rime, B.K.
J. Gen. Virol. 73, 1189-1194, 1992
A:Title: Molecular characterization of phocine distemper virus: Gene order and sequen
A:Reference number: J01535; MUID:92268877; PMID:1588321
A:Accession: J01535
A:Molecule type: mRNA
A:Residues: 1-607 <CNR>
A:Cross-references: GB:D10371
C:Genetics:
A:Gene: H
C:Superfamily: measles virus hemagglutinin
C:Keywords: glycoprotein; hemagglutinin; transmembrane protein
F:39-55/Domain: transmembrane #status predicted <TMN>
F:19,149,276,391,422,456,587/Binding site: carbohydrate (asn) (covalent) #status pred

Query Match 31.9% Score 44; DB 1; Length 607;
Best Local Similarity 28.6% Pred. No. 9.4;
Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

OY 3 IATLCTKXXXXVLDGIDKXXXXLIQLI 30
DB 292 LASLCTKESTILNLGDESONSVLVI 319

RESULT 7

T32510
hypothetical protein C44B12.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 04-Mar-2000
C:Accession: T32510
R:tin-mollam, A.
Submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid C44B12.
A:Reference number: 221183
A:Accession: T32510
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-246 <TIN>
A:Cross-references: EMBL:AF036692; PIDN:AA88326.1; GSPDB:GN00022; CESP:C44B12.3
C:Genetics:
A:Gene: CESP:C44B12.3
A:Map position: 4
A:Introns: 73/3; 112/3; 137/1; 202/1
C:Superfamily: Caenorhabditis elegans hypothetical protein C44B12.3

Query Match 31.2%; Score 43; DB 2; Length 246;
Best Local Similarity 45.0%; Pred. No. 5.6;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 14 VDFGIDKXXXLIQIEDK 33
DB 145 MIFGIDKWKWQMSIGIEEK 164

RESULT 8
T18111
probable UDPglucose 6-dehydrogenase (EC 1.1.1.22) - Chlorella virus PCV-1
C:Species: Chlorella virus PCV-1
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18111
R:Graves, M.V.; Van Etten, J.L.
Submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: T18111
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-389 <GRA>
A:Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96945.1
A:Experimental source: specific host Chlorella strain NC64A
C:Genetics:
A:Note: A609L
C:Superfamily: GDPmannose dehydrogenase
C:Keywords: oxidoreductase

Query Match 31.2%; Score 43; DB 2; Length 389;
Best Local Similarity 31.6%; Pred. No. 9;
Matches 12; Conservative 6; Mismatches 10; Indels 10; Gaps 2;

QY 1 GFATLCT---KXXXVLDGIDKXXXLIQIEDK 33
DB 10 GYVGTACAVLAAKNEIVLIDISDR-----VOLIKNK 42

RESULT 9
A53455
syntaxin-binding protein n-Sect - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Aug-1995 #sequence_revision 25-Aug-1995 #text_change 05-Nov-1999
C:Accession: A53455; A53487
R:Peversner, J.; Hsu, S.C.; Scheller, R.H.
Proc. Natl. Acad. Sci. U.S.A. 91, 1445-1449, 1994
A:Title: n-Sect: a neural-specific syntaxin-binding protein.
A:Reference number: A53455; MUID:94151347; PMID:8108429
A:Accession: A53455
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-594 <PEV>
A:Cross-references: GB:U26264; NID:g459948; PIDN:AA19246.1; PID:g459949
R:Garcia, E.P.; Gatti, E.; Butler, M.; Burton, J.; De Camilli, P.

Proc. Natl. Acad. Sci. U.S.A. 91, 2003-2007, 1994
A:Title: A rat brain Sect homologue related to Rop and UNC18 interacts with syntaxin.
A:Reference number: A53487; MUID:94181519; PMID:8134339
A:Accession: A53487
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-594 <GAR>
A:Cross-references: GB:U06069; NID:g458341; PIDN:AA17987.1; PID:g458342
C:Keywords: brain; membrane trafficking

Query Match 31.2%; Score 43; DB 2; Length 594;
Best Local Similarity 45.5%; Pred. No. 14;
Matches 15; Conservative 1; Mismatches 15; Indels 2; Gaps 1;

QY 3 IATLCT--KXXXVLDGIDKXXXLIQIEDK 33
DB 176 IATLCATLKEYPAVRVRYGKYNALLAQIIDK 208

RESULT 10
H83624
probable porin PA0162 (imported) - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Jun-2001
C:Accession: H83624
R:Stoyer, C.R.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83624
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-444 <STO>
A:Cross-references: GB:AE004454; GB:AE004091; NID:g9945990; PIDN:AAG03552.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0162
C:Superfamily: Pseudomonas aeruginosa outer membrane porin PA0958

Query Match 30.4%; Score 42; DB 2; Length 444;
Best Local Similarity 45.0%; Pred. No. 16;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 GFATLCTKXXXVLDGID 20
DB 64 GFATFQSGYTPGVGFGVD 83

RESULT 11
C64476
hypothetical protein MJ1412 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 13-Sep-1996
C:Accession: C64476
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Meinkoth, K.G.; Merrick, J.M.; Glodex, R.; rson, J.D.; Sadow, P.W.; Hann, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: C64476
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-498 <BUL>
A:Cross-references: GB:U67581; GB:L77117; NID:g1592052; PID:g1592061; TIGR:MJ1412; PI
C:Genetics:
A:Map position: FOR137212-1374708
C:Superfamily: hypothetical protein MJ1412

Query Match 30.4%; Score 42; DB 2; Length 498;

	Best Local Similarity	50.0%	Pred. No. 18:			
	Matches	8;	Conservative	3;	Mismatches	5;
						Indels
						0;
						Gaps
						0.
Qy	18	GIDKXXXLIQLIEDK	33			
				:		
Db	363	GIDKNNIIVIELFEDK	378			

RESULT 12
F83550

```
serine proteinase MucB precursor PA0766 | imported | - pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
```

A;Gene: mucD; PA0766
C;Superfamily: Helicobacter serine proteinase

```

      11:11      |:11:11
Db 287 EGLDKPSGALVAQLVED 303

```

RESULT_13
H90208
hypothetical protein pyrF [imported] - *Sulfolobus solfataricus*
8, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 8

Query Match	29.7%	Score 41:	DB 2:	Length 222:
Best Local Similarity	56.2%	Pred No. 12:		
Matches	9:	Conservative	1:	Mismatches 6: Indels 0: Gaps 0

RESULT 14
D42510
OIL protein - vaccinia virus (strain Copenhagen)
C:Species: vaccinia virus

OIL protein - vaccinia virus (strain Copenhagen),
C:Species: vaccinia virus

A:Note: host Homo sapiens (man)
C:Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 20-Jun-20000
C:Accession: D42510
R:Johnson, G. P.
Submitted to GenBank, June 1990
A:Reference number: A33172

C: Superfamily: variola major virus hypothetical protein Q1L

RESULT 15
T28491

hypothetical protein QIL - variola major virus

Query Match	29.7%	Score 41	DB 2	length 666
Best Local Similarly	31.0%	Pred. No.	36	
Matches 9	Conservative 6	Mismatches 14	Indels 0	Gaps 0

Search completed: May 9, 2003, 16:18:44
Job time : 26.2157 secs

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce>)

```

CC or send an email to license@lsb-sib.ch).
CC
DR EMBL: M83945: AAA29111.1: -
DR EMBL: X70851: CAA50203.1: -
DR PIR: S25283: S25283.
DR PIR: A41279: A41279.
DR InterPro: IPR000004: SAPB.
DR SMART: SM00118: SAPB: 1.
KW Signal: Antibiotic.
FT SIGNAL 1 21
FT PEPTIDE 22 98
FT DISULFID 26 98
FT DISULFID 29 92
FT DISULFID 56 67
FT VARIANT 71 71
SO SEQUENCE 98 AA: 10504 MW: F9592FDCIFACFCFFA CRC64:

Query Match 73.9% Score 102: DB 1: Length 98:
Best Local Similarity 75.8% Pred. No. 7.7e-12:
Matches 25: Conservative 0: Mismatches 0: Indels 8: Gaps 2:

Oy 1 GFIAFLCTKXXXXVDFGIDKXXXXLLIQLIDDK 33
Db 61 GFIAFLCTK---VDFGIDK---LIQLIDDK 85

RESULT 2
PPNP_ENTH1
ID PPNP_ENTH1 STANDARD: PRT: 97 AA.
AC 007831:
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nonpathogenic pore-forming peptide precursor (APNP).
DE Entamoeba histolytica.
CC Eukaryote: Entamoebidae; Entamoeba.
OX NCBI_TaxID=5759;
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-34.
RC STRAIN-SAW 142;
RX MEDLINE=93295428; PubMed=8515772;
RA Lelepe M., Bahr E., Tannich E., Horstmann R.D.;
RT "Comparison of pore-forming peptides from pathogenic and
RT nonpathogenic Entamoeba histolytica.";
RL Mol. Biochem. Parasitol. 59:101-110(1993).
CC -1- FUNCTION: FORMS PORE IN THE MEMBRANE OF HOST TISSUES. IMPLICATED
CC IN THE CYTOLYTIC ACTIVITY OF THE PARASITE.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: CONTAINS 1 SAOSIN B-TYPE DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/
CC or send an email to license@lsb-sib.ch).
CC
CC
CC EMBL: L04984: AAA18632.1: -
DR InterPro: IPR000004: SAPB.
DR SMART: SM00118: SAPB: 1.
KW Signal.
FT SIGNAL 1 20
FT CHAIN 21 97
FT DISULFID 25 97
FT DISULFID 28 91
FT DISULFID 55 66
SO SEQUENCE 97 AA: 10432 MW: AE171C58B86EED07 CRC64:

Query Match 73.2% Score 101: DB 1: Length 97:
Best Local Similarity 72.7% Pred. No. 1.2e-11:
Matches 24: Conservative 1: Indels 8: Gaps 2:

```

```

Oy 1 GFIAFLCTKXXXXXVDFGIDKXXXXXLIOLIEDK 33
    |||||
Db 60 GFIAFLCTK-----VLDPGVDK-----LIOLIEDK 84

RESULT 3
PEPB_ENTH1
ID PEPB_ENTH1 STANDARD: PRT: 96 AA.
AC 024824;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pore-forming peptide ameobapore B precursor (EH-APP).
OS Entamoeba histolytica.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-63.
RC STRAIN=HM-1;IMSS:
RX MEDLINE=955231296; PubMed=7715451;
RA Leippe M., Andrae J., Nickel R., Tannich E., Mueller-Eberhard H.J.:
RT "Ameobapores, a family of membranolytic peptides from cytoplasmic
RT granules of Entamoeba histolytica: isolation, primary structure, and
RT pore formation in bacterial cytoplasmic membranes."
RL Mol. Microbiol. 14:895-904(1994).
CC -1- FUNCTION: FORMS PORES IN THE CYTOPLASMIC MEMBRANE OF HOST CELLS.
CC HAS ANTIBACTERIAL ACTIVITY AGAINST M.LUTEUS, NO ACTIVITY AGAINST
CC E.COLI. IMPLICATED IN THE CYTOLYTIC ACTIVITY OF THE PARASITE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES.
CC -1- SIMILARITY: CONTAINS 1 SAPOSIN B-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC
CC DR EMBL: X76904; CAA54226.1; -.
CC DR InterPro: IPR000004; SAPB.
CC DR SMART: SM00118; SAPB; 1.
CC KW SIGNAL; Antibiotic.
FT SIGNAL 1 19
FT PEPTIDE 20 96 PORE-FORMING PEPTIDE AMEOBAPORE B.
FT DISULFID 24 96 BY SIMILARITY.
FT DISULFID 27 90 BY SIMILARITY.
FT DISULFID 54 65 BY SIMILARITY.
SQ SEQUENCE 96 AA: 10437 MW: 9956AFADBD02B42CC CRC64:
Query Match 45.7%; Score 63; DB 1; Length 96;
Best Local Similarity 43.8%; Pred. No. 0.0013;
Matches 14; Conservative 7; Indels 8; Gaps 2;
Oy 1 GFIAFLCTKXXXXXVLDPGIDKXXXXXLIOLIED 32
    ||:||||:|:||||:|:||||:
Db 59 GFLGLTCEK-----ILSFGVDE-----LVKLLEN 82

RESULT 4
PEPC_ENTH1
ID PEPC_ENTH1 STANDARD: PRT: 101 AA.
AC 024825;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pore-forming peptide ameobapore C precursor (EH-APP).
OS Entamoeba histolytica.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-69.
RC STRAIN=HM-1;IMSS:

```


RX MEDLINE-95231296; PubMed-7715451.
 RA Leippe M., Andree J., Nickel R., Tannich E., Mueller-Eberhard H.J.;
 RT "Amoebapores", a family of membranolytic peptides from cytoplasmic
 RT granules of Entamoeba histolytica: isolation, primary structure, and
 RT pore formation in bacterial cytoplasmic membranes.";
 RL Mol. Microbiol. 14:895-904(1994).
 CC -1- FUNCTION: FORMS PORES IN THE CYTOPLASMIC MEMBRANE OF HOST CELLS.
 CC HAS ANTIBACTERIAL ACTIVITY AGAINST M. LUTERUS, NO ACTIVITY AGAINST
 CC E. COLI. IMPLICATED IN THE CYTOLYTIC ACTIVITY OF THE PARASITE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES.
 CC -1- SIMILARITY: CONTAINS 1 SAPOSIN B-TYPE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X76904; CAA54225.1; .
 DR InterPro: IPR000004; SAPP.
 DR SMART: SM00118; SAPP; 1.
 KW Signal; Antidiabetic.
 FT SIGNAL 1 24
 FT PEPTIDE 25 101 PORE-FORMING PEPTIDE AMOEBOPORE C.
 FT DISULFID 29 101 BY SIMILARITY.
 FT DISULFID 32 95 BY SIMILARITY.
 FT DISULFID 59 70 BY SIMILARITY.
 SQ SEQUENCE 101 AA; 10855 MW; 28AC72135E73499B CRC64;
 Query Match 39.1%; Score 54; DB 1; Length 101;
 Best Local Similarity 35.7%; Pred. No. 0.0064;
 Matches 10; Conservative 7; Mismatches 7; Indels 4; Gaps 1;
 Oy 1 GFATLCTKXXXXVDFGIDKXXXXLQ 28
 1 : ||||| : : ||||| : :
 Db 64 GLVERLCTK---IVSYGIDKLEIKLE 87
 RESULT 5
 STBL_HUMAN STANDARD; PRT; 594 AA.
 ID STBL_HUMAN 064320; 062759; 028208;
 AC 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Syntaxin binding protein 1 (Unc-18 homolog) (Unc-18A) (Unc-18-1) (N-
 DE Sec1) (rbsec1) (p67).
 GN STXBPL OR UNC18A.
 OS Homo sapiens (Human).
 OS Rattus norvegicus (Rat), and
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606, 10116, 9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human; TISSUE-Brain;
 RA MEDLINE-96421662; PubMed-8824310;
 RA Gengyo-Ando K., Kitayama H., Mukaida M., Ikawa Y.;
 RT "A murine neural-specific homolog corrects cholinergic defects in
 RT Caenorhabditis elegans unc-18 mutants.";
 RL J. Neurosci. 16:6695-6702(1996).
 RN [2]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC SPECIES-Human;
 RA MEDLINE-98207254; PubMed-9545644;
 RA Swanson D.A., Steel J.M., Valle D.;
 RT "Identification and characterization of the human ortholog of rat
 RT STXBPL, a protein implicated in vesicle trafficking and
 RT neurotransmitter release.";
 RL Genomics 48:373-376(1998).

RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Human; TISSUE-Skin;
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC SPECIES-Rat, and Bovine; TISSUE-Brain;
 RX MEDLINE-94067319; PubMed-8247129;
 RA Hata Y., Slaughter C.A., Suedhof T.C.;
 RT "Synaptic vesicle fusion complex contains unc-18 homologue bound to
 RT syntaxin.";
 RL Nature 366:347-351(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Rat; TISSUE-Brain cortex;
 RX MEDLINE-94151347; PubMed-6108429;
 RA Pevsner J., Hsu S.-C., Scheller R.H.;
 RT "n-Sec1, a neural-specific syntaxin-binding protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:1445-1449(1994).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Rat; TISSUE-Brain;
 RX MEDLINE-94181519; PubMed-8134339;
 RA Garcia E.P., Gatti E., Butler M., Burton J., de Camilli P.;
 RT "A rat brain Sec1 homologue related to KOP and UNC18 interacts with
 RT syntaxin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:2003-2007(1994).
 RN [7]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RC SPECIES-Rat; STRAIN-Sprague-Dawley; TISSUE-Brain;
 RX MEDLINE-95213277; PubMed-7698978;
 RA Garcia E.P., McPherson P.S., Chilcote T.J., Takei K., de Camilli P.;
 RT "n-Sec1 and B colocalize with syntaxin 1 and SNAP-25 throughout the
 RT axon, but are not in a stable complex with syntaxin.";
 RL J. Cell Biol. 129:105-120(1995).
 RN [8]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC SPECIES-Rat; TISSUE-Brain;
 RX MEDLINE-95239157; PubMed-7536802;
 RA Shetty K.T., Kaech S., Link W.T., Jaffe H., Flores C.M., Wray S.,
 RA Pant H.C., Beushausen S.;
 RT "Molecular characterization of a neuronal-specific protein that
 RT stimulates the activity of Cdk5.";
 RL J. Neurochem. 64:1988-1995(1995).
 RN [9]
 RP SEQUENCE FROM N.A.
 RC SPECIES-Bovine; TISSUE-Adrenal medulla;
 RA Wang G.-T., Gillis K.D.;
 RT "Bovine chromaffin cell munc18-1 cDNA deduced amino acid sequence is
 RT identical to that of rat munc18-1.";
 RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: MAY PARTICIPATE IN THE REGULATION OF SYNAPTIC VESICLE
 CC DOCKING AND FUSION, POSSIBLY THROUGH INTERACTION WITH GTP-BINDING
 CC PROTEINS. ESSENTIAL FOR NEUROTRANSMISSION AND BINDS SYNTAXIN, A
 CC COMPONENT OF THE SYNAPTIC VESICLE FUSION MACHINERY PROBABLY IN A
 CC 1:1 RATIO. CAN INTERACT WITH SYNTAXINS 1, 2, AND 3 BUT NOT
 CC SYNTAXIN 4. MAY PLAY A ROLE IN DETERMINING THE SPECIFICITY OF
 CC INTRACELLULAR FUSION REACTIONS.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC AND MEMBRANE-ASSOCIATED.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1/A (shown here) and 2/BE; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: BRAIN AND SPINAL CORD. HIGHLY ENRICHED IN
 CC AXONS.
 CC -1- DEVELOPMENTAL STAGE: PAINT LEVELS ARE DETECTABLE AT EMBRYONIC DAY
 CC 14, WITH LEVELS RISING AT LATER EMBRYONIC AGES AND PEAKING AT
 CC POSTNATAL DAY 7.
 CC -1- SIMILARITY: BELONGS TO THE STXBPL/UNC-18/SEC1 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

```
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
-----
DR EMBL; D63851; BAA19483.1; -
DR EMBL; AF004562; AAC39688.1; -
DR EMBL; BC015749; AAI15749.1; -
DR EMBL; L26087; AAB41113.1; -
DR EMBL; L26088; AAB41112.1; -
DR EMBL; L26264; AAB19246.1; -
DR EMBL; U00609; AAA17987.1; -
DR EMBL; U21116; AAA69350.1; -
DR EMBL; AF153327; AAD37018.1; -
DR Genew; HGNC:11444; STXBPI.
DR MIM; 602926; -
DR InterPro; IPR001619; Sec1-1like.
DR Pfam; PF00995; Sec1_1.
KW Protein transport; Alternative splicing.
FT VARSPLIC 576 594 OKLDITKKLKTKDEETSS -> TKFLMDLRHHPDRESSRV
FT FT SFEQAPRME (IN ISOFORM 2).
FT FT L -> V (IN REF. 4; AAB41112).
SQ SEQUENCE 594 AA; 67568 MW; 2DD0715F875CE0F3 CRC64;
Query Match 31.2%; Score 43; DB 1; Length 594;
Best Local Similarity 45.5%; Pred. No. 4.2;
Matches 15; Conservative 1; Mismatches 15; Indels 2; Gaps 1;
OY 3 IATVLC--KXXXXVLDFGIDKXXXLLQLIEDK 33
DB 176 IATCATLKEYPVAVRVREGEYKDMLLAQLDIQOK 208
IIIII I I I I IIII
RESULT 6
STBI_MOUSE STANDARD; PRT; 594 AA.
AC 008599:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE SynTaxin binding protein 1 (Unc-18 homolog) (Unc-18A) (Unc-18-1).
CN STXBPI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=96421662; PubMed=8824310;
RA Gengyo-Ando K., Kitayama H., Mukaida M., Itawa Y.;
RA "A murine neural-specific homolog corrects cholinergic defects in
RA Caenorhabditis elegans unc-18 mutants.";
RL J. Neurosci. 16:6695-6702(1996).
CC -1- FUNCTION: MAY PARTICIPATE IN THE REGULATION OF SYNAPTIC VESICLE
CC DOCKING AND FUSION, POSSIBLY THROUGH INTERACTION WITH GTP-BINDING
CC PROTEINS. ESSENTIAL FOR NEUROTRANSMISSION AND BINDS SYNTAXIN, A
CC COMPONENT OF THE SYNAPTIC VESICLE FUSION MACHINERY PROBABLY IN A
CC 1:1 RATIO. CAN INTERACT WITH SYNTAXINS 1, 2, AND 3 BUT NOT
CC SYNTAXIN 4. MAY PLAY A ROLE IN DETERMINING THE SPECIFICITY OF
CC INTRACELLULAR FUSION REACTIONS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE STXBP/UNC-18/SEC1 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
-----
DR EMBL; DA5903; BAA19479.1; -
DR MCD; MG1:107363; Stxbpl.
DR InterPro; IPR001619; Sec1-like.
```

```

DR Pfam: PF00995; Scl: 1.
KW Protein transport.
SQ SEQUENCE 594 AA; 67701 MW; 63E7628E15B3E847 CRC64;

Query Match
Best Local Similarity 31.2%; Score 43; DB 1; Length 594;
Matches 15; Conservative 1; Mismatches 15; Indels 2; Gaps 1;

OY 3 IATLCT--KXXXXVLDPCGIDKXXXXLQIILEDK 33
      ||||| | | | | | | | | | | | | | | |
DB 176 IATLCATLKEYPAYRYRGCEYKDNALLAQIQDK 208

RESULT 7
STA_METKA
ID SVA_METKA STANDARD: PRT: 915 AA.
AC 08TWY1;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine-tRNA ligase) (Alars).
GN ALAS OR MK0900.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyrataceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN (1)
RP SEQUENCE FROM N.S.A.
RC STRAIN-AV19 / DSM 6324 / JCM 9639;
RX MEDLINE=21927647; PubMed=11930014;
RA Slegarev A.I., Mezheva K.V., Makrova K.S., Polushin N.N.,
RA Shchekardinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA Maljch A.G., Koonin E.V., Kozayvkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and morphology of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4664-4669(2002).
CC -I- CATALYTIC ACTIVITY: ATP + L-alanine + CRNA(Ala) = AMP +
CC diphosphate + L-alanyl-tRNA(Ala).
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: AE010379; AAM02113.1; ALT_INIT.
CC DR PROSITE: PS50860; AA-TRNA_LIGASE_I1_ALA.1.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC SQ SEQUENCE 915 AA; 103908 MW; 6DD88BA193C84E CRC64;

Query Match
Best Local Similarity 30.1%; Score 41.5; DB 1; Length 915;
Matches 9; Conservative 29.0%; Pred. No. 12;
Matches 9; Conservative 9; Mismatches 12; Indels 1; Gaps 1;

OY 1 GFIAITLCTKXXXXVLDPCGIDKXXXXLQIILEDK 31
      ||||| | | | | | | | | | | | | | | |
DB 355 GYLARLVIRALRLILD-GIDAREYRELLVEVE 384

RESULT 8
PYRF_SULSO
ID PYRF_SULSO STANDARD: PRT: 222 AA.
AC 09UX10;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)

```

DE (OMPDCase) (OMPDCase).
GN PYR OR SSO0616 OR C08_034.
OS Sulfolobus solfataricus.
OC Archaea: Crenarchaeota: Thermoprotei: Sulfolobales: Sulfolobaceae;
CC Sulfolobus.
NCBI_TaxID=2287;
RN
RP
RX MEDLINE-ATCC 35092 / DSM 1617 / P2;
MEDLINE-20165948; PubMed-10701121;
RA Charlebois R.L., Singh R.K., Chan-Weher C.C.-Y., Allard G., Chow C.,
Confalonieri F., Curtis B., Duguet M., Erasus G., Faquy D.,
Gaasterland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
Kushwana N., Lafleur E., Medina N., Peng X., Penny S.L., She Q.,
St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
Ragan M.A., Sengen C.W.;
RT "Gene content and organization of a 281-kbp contig from the genome of
the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
RL Genome 43:116-136(2000).
RN
RP
RX MEDLINE-ATCC 35092 / DSM 1617 / P2;
MEDLINE-21332296; PubMed-11427726;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
Aweyer M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
De Moors A., Erasus G., Fletcher C., Gordon P.M.K.,
Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
Thi-NGOC H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
Garrett R.A., Ragan M.A., Sengen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate - UMP + CO(2).
CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.
CC -1- SUBUNIT: Homodimer (by similarity).
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY. SUBFAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y18930; CAB57684.1; -
DR EMBL: AF006590; AAK40927.1; -
DR HSPB: 026232; IDV7.
DR InterPro: IPR001754; OMPDecase.
DR Pfam: PF00215; OMPDecase; 1.
DR PROSITE: PS00156; OMPDecase; FALSE_NEG.
KW Lyase: Decarboxylase; Pyrimidine biosynthesis; Complete proteome.
FT ACT SITE 61
FT PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 222 AA; 24701 MW; 48F3E83DB0F033FF CRC64;
Query Match 29.7%; Score 41; DB 1; Length 222;
Best Local Similarity 56.2%; Pred. No. 3.6;
Matches 9; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
OY 14 VLDGIDKXXXXLQIOL 29
DB 36 VLDLGVDTRELLIGL 51
RESULT 9
VOOL_VACCC STANDARD: PRT; 666 AA.
AC P21093;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein O1.
GN O1L.

OS Vaccinia virus (strain Copenhagen).
OC Viruses: dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OC
NCBI_TaxID=10249;
RN
RP
RX MEDLINE-91021027; PubMed-2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
Paoletti E.;
RT "The complete DNA sequence of vaccinia virus.";
Virology 179:247-266(1990).
RL
RN
RP COMPLETE GENOME.
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
Paoletti E.;
RT "Appendix to 'The complete DNA sequence of vaccinia virus'.";
Virology 179:517-563(1990).
CC -1- SIMILARITY: BELONGS TO THE POXVIRUSES O1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M35027; AAA48053.1; -
DR PIR: D42510; D42510.
SQ SEQUENCE 666 AA; 77577 MW; 46DF3299DDEB66EF CRC64;
Query Match 29.7%; Score 41; DB 1; Length 666;
Best Local Similarity 31.0%; Pred. No. 11;
Matches 9; Conservative 6; Mismatches 14; Indels 0; Gaps 0;
OY 4 ATLCGRXXXXVLDGIDKXXXXLQIOLIED 32
DB 129 ATYLLKINAYMINFKRIDLVDEIDLVKD 157
RESULT 10
VOOL_VARY STANDARD: PRT; 666 AA.
AC P34010;
ID VOOL_VARY
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Protein O1.
GN O1L OR Q1L.
OS Variola virus.
OC Viruses: dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OC
NCBI_TaxID=10255;
RN
RP
RX MEDLINE-91021027; PubMed-2219722;
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
Paoletti E.;
RT "The complete DNA sequence of vaccinia virus.";
Virology 179:247-266(1990).
RL
RN
RP COMPLETE GENOME.
RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
Paoletti E.;
RT "Appendix to 'The complete DNA sequence of vaccinia virus'.";
Virology 179:517-563(1990).
CC -1- SIMILARITY: BELONGS TO THE POXVIRUSES O1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X69198; CAA48994.1; -
 DR PIR: F36842; F36842.
 SQ SEQUENCE 666 AA; 77305 MW; F9E9192C93A6BAAF CRC64;

Query Match 29.7% Score 41; DB 1; Length 666;
 Best Local Similarity 31.0%; Pred. No. 11;
 Matches 9; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

OY 4 TLCTKXXXXVDFGIDKXXXXLIQIED 32
 DB 129 ATYLKINAMINFKIDLIVDEIIDLVD 157

RESULT 11

ALAB_ARATH STANDARD; PRT: 1189 AA.

AC 09LX90: 2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Potential phospholipid-transferring ATPase 8 (EC 3.6.3.1)
 DE (Amphophospholipid flippase 8).
 GN ALAB OR AT3G2870 OR K16N12.9.

OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsids.
 RN NCBI_TaxID=3702;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=20363099; PubMed=10907853;
 RA Kaneko T., Katoh T., Sato S., Nakamura A., Asanishi E., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the 4,251,695 bp regions covered by 90 pI, TAC
 RT and BAC clones."

CC -1- FUNCTION: INVOLVED IN TRANSPORT OF PHOSPHOLIPIDS (POTENTIAL).
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O -> ADP + phosphate.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY
 CC (E1-E2 ATPASES). SUBFAMILY IV.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AP000371; BAB02533.1; -
 DR InterPro: IPR001757; ATPase_E1-E2.
 DR InterPro: IPR001454; Hlgase/hydrase.

DR Pfam: PF00702; Hydrolase: 1.
 DR PRINTS: PR00119; CATATPASE.
 DR PROSITE: PS00154; ATPASE_E1-E2: 1.
 KM Hydrolyase; Transmembrane; Phosphorylation; ATP-binding; Metal-binding;
 KM Magnesium; Multigene family.

FT DOMAIN 1 71 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 72 93 POTENTIAL.
 FT DOMAIN 94 97 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 98 120 POTENTIAL.
 FT DOMAIN 121 299 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 300 321 POTENTIAL.
 FT DOMAIN 322 358 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 359 376 POTENTIAL.
 FT DOMAIN 377 920 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 921 940 POTENTIAL.
 FT DOMAIN 941 954 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 955 974 POTENTIAL.
 FT DOMAIN 975 1004 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 1005 1027 POTENTIAL.
 FT DOMAIN 1028 1040 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1041 1063 POTENTIAL.
 FT DOMAIN 1064 1090 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 1070 1090 POTENTIAL.
 FT DOMAIN 1091 1107 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1108 1132 POTENTIAL.
 FT DOMAIN 1133 1189 CYTOPLASMIC (POTENTIAL).
 FT MOD_RES 1189 1189 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 424 424 MAGNESIUM (BY SIMILARITY).
 FT METAL 865 865 MAGNESIUM (BY SIMILARITY).
 FT METAL 869 869 MAGNESIUM (BY SIMILARITY).
 SQ SEQUENCE 1189 AA; 135309 MW; DCF951CF3FC55E83 CRC64;

Query Match 29.7% Score 41; DB 1; Length 1189;
 Best Local Similarity 36.4%; Pred. No. 20;
 Matches 12; Conservative 4; Mismatches 13; Indels 4; Gaps 1;

OY 5 TLCTKXXXXVDFGIDKXXXXLIQI-----TEPK 33
 DB 670 TLVTEDRDALIDMAADKIEKDLILGSTAVEDK 702

RESULT 12

TP2B_MOUSE STANDARD; PRT: 1612 AA.

AC 064511;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA topoisomerase II, beta isozyme (EC 5.99.1.3).
 GN TOP2B.

OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCBI_TaxID=10090;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Brain;
 RA Miyake M., Adachi N., Kikuchi A.;
 RT Submitted (SEP-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CONTROL OF TOPOLOGICAL STATES OF DNA BY TRANSIENT
 CC BREAKAGE AND SUBSEQUENT REJOINING OF DNA STRANDS. TOPOISOMERASE II
 CC MAKES DOUBLE-STRAND BREAKS.
 CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
 CC of double-stranded DNA.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear; nucleolar.
 CC -1- MISCELLANEOUS: EUKARYOTIC TOPOISOMERASE I AND II CAN RELAX BOTH
 CC NEGATIVE AND POSITIVE SUPERCOILS, WHEREAS PROKARYOTIC ENZYMES
 CC RELAX ONLY NEGATIVE SUPERCOILS.

CC -1- SIMILARITY: BELONGS TO THE TYPE II TOPOISOMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D38046; BAA07236.1; -
 DR HSSP: P06786; IBCM.

DR MGD: MGI:98791; Top2b.
 DR InterPro: IPR003594; ATPbind ATPase.
 DR InterPro: IPR003857; CBFA_NFYB_topis.
 DR InterPro: IPR001241; DNA_topoisot1.
 DR InterPro: IPR002205; DNA_topoisot1v.


```

RESULT 15
VG33_HSV11 STANDARD: PRT: 290 AA.
AC 000118;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Hypothetical gene 33 protein.
GN 33.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Ictalurid Herpes-like viruses.
OX NCBI_TaxID=10401;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=Aburn 1;
RX MEDLINE=92087490; PubMed=1727613;
RA Davison A.J.;
RT "Channel catfish virus: a new type of herpesvirus.";
RL Virology 186:9-14(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M75136; AAA88136.1;
DR PIR: G36789; G36789.
KW Hypothetical protein.
SQ SEQUENCE 290 AA: 34173 MW: 21C26260985A7E5E CRC64:

```

```

Query Match 28.3%; Score 39; DB 1: Length 290;
Best Local Similarity 77.8%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 GFATLCTK 9
||| |||
DB 217 GFITTTCTK 225

```

Search completed: May 9, 2003, 16:16:16
 Job time : 14.9706 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - protein search, using sw model

Run on: May 9, 2003, 16:10:40 / Search time 25.3922 Seconds
(without alignments)
300.240 Million cell updates/sec

Title: US-09-851-422a-1
Perfect score: 138
Sequence: 1 GFATLCTKXXXXVLDGIDKXXXXXLIQLIEDKXXXX 37

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64	46.4	96	5	Q9U8G5
2	49	35.5	101	5	Q9U8G4
3	46	33.3	87	5	O15591
4	44	31.9	215	16	Q9PHP9
5	44	31.9	600	10	Q9LS72
6	44	31.9	607	12	Q84690
7	44	31.9	607	12	O56986
8	43.5	31.5	388	16	Q8XN50
9	43.5	31.5	807	16	Q9H5D7
10	43	31.2	162	5	O44142
11	43	31.2	389	12	O41091
12	43	31.2	431	11	Q9DBR9
13	43	31.2	594	11	O88290
14	43	31.2	594	11	O8VD51
15	43	31.2	603	4	Q96TG8
16	42	30.4	241	13	O57532

17	42	30.4	444	16	Q916X0
18	42	30.4	498	17	O58807
19	41.5	30.1	474	2	O9AL51
20	41.5	30.1	474	16	O57155
21	41.5	30.1	918	17	O8TWI1
22	41	29.7	254	12	O8V2V2
23	41	29.7	330	5	O8QO58
24	41	29.7	330	5	O9VP58
25	41	29.7	388	10	O8W330
26	41	29.7	665	12	O8V520
27	41	29.7	666	12	O8S376
28	41	29.7	666	12	O9JFC9
29	41	29.7	666	12	O9QJNC
30	41	29.7	666	12	O9PXS8
31	41	29.7	666	12	O8QXW0
32	41	29.7	1094	10	O22220
33	40.5	29.3	674	5	O09403
34	40	29.0	149	10	O8S4R6
35	40	29.0	217	16	O67058
36	40	29.0	309	10	O9AX19
37	40	29.0	487	16	O97579
38	40	29.0	754	10	O9S8R5
39	40	29.0	985	17	O97AX8
40	40	29.0	1007	4	O43379
41	39	28.3	149	10	O8S4R5
42	39	28.3	213	17	O978A2
43	39	28.3	223	16	O9JRP1
44	39	28.3	249	5	O9U491
45	39	28.3	253	16	O9K3V7

ALIGNMENTS

RESULT 1

ID Q9U8G5 PRELIMINARY: PRT: 96 AA.
AC Q9U8G5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pore-forming protein isoform B precursor.
DE DP-B.
OS Entamoeba dispar.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=46681;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SAW 142;
RX MEDLINE=99449604; Pubmed=10518795;
RA Nickel R., Ott C., Dandekar T., Leippe M.,
RT to amoebapores in structure, expression and divergence
RT Eur. J. Biochem. 265:1002-1007(1999).
RL EMBL: AF082528; AAF04195.1; --
DR InterPro: IPR000004; SApB.
DR ProDom: PDOM01732; SApB; 1.
DR SMART; SM00118; SApB; 1.
KW SIGNAL.
FT CHAIN 1 19 POTENTIAL.
FT CHAIN 20 96 PORE-FORMING PROTEIN ISOFORM B.
SQ SEQUENCE 96 AA: 10480 MW: 9FA4ZAC9A9FARB2 CRC64;

Query Match 46.4%; Score 64; DB 5; Length 96;
Best Local Similarity 43.8%; Pred. No. 0.0003;
Matches 14; Conservative 7; Mismatches 3; Indels 8; Gaps 2;

QY 1 GFATLCTKXXXXVLDGIDKXXXXXLIQLIED 32
DB 59 GFLGTLCKN-----ILSGVDE-----LVKLIEN 82

RESULT 2

ID	Q9UBG4	PRELIMINARY:	PRC:	101 AA.
AC	Q9UBG4.			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Pore-forming protein Isoform C precursor.			
CN	DP-C.			
OS	Entamoeba dispar.			
OC	Eukaryota; Entamoebidae; Entamoeba.			
OX	NCBI_TaxID=45681;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SAW 142;			
RX	MEDLINE=99449604; PubMed=10518795;			
RA	Nickel R., Olt C., Dandekar T., Leippe M.;			
RT	"Pore-forming peptides of entamoeba dispar similarity and divergence			
RT	to amoebozoetes in structure, expression and activity.";			
RL	Eur. J. Biochem. 265:1002-1007(1999).			
DR	EMBL; AF082529; AAF04196.1; -.			
DR	InterPro; IPR000004; SApB.			
DR	Prodom; PD001732; SApB; 1.			
DR	SMART; SM00118; SApB; 1.			
KW	Signal.			
FT	SIGNAL. 1 24			POTENTIAL.
FT	CHAIN 25 101			PORE-FORMING PROTEIN ISOFORM C.
SO	SEQUENCE 101 AA; 10877 MW; 7699A7EA9E0876EB CRC64;			

Query Match	35.5%	Score 49	DB 5	Length 101
Best Local Similarity	35.7%	Pred. No. 0.22		
Matches	10	Conservative	7	Mismatches 7; Indels 4; Gaps 1
QY	1	GFIAFLCTKXXXXVLPDGIKXXXXXLIQ	28	
DB	64	GVIEFVCSK---IVSYGIDKLEIKETI	87	

RESULT 3			
015591			
ID	015591	PRELIMINARY;	PRT; 87 AA.
AC	015591;		
DT	01-JAN-1998 (TREMBLrel. 05, Created)		
DT	01-JAN-1998 (TREMBLrel. 05, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	Amcebapore C homologue (Fragment).		
OS	Entamoeba histolytica.		
OC	Eukaryota; Entamoebidae; Entamoeba.		
OX	NCBI_TaxID=5759;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HM-1;		
RX	MEDLINE=97396144; PubMed=9245698;		
RA	Tanaka T., Tanaka M., Mitsui Y.;		
RT	"Analysis of expressed sequence tags (ESTs) of the parasitic protozoa		
RT	Entamoeba histolytica.";		
RL	Biochem. Biophys. Res. Commun. 236:611-615(1997).		
DR	EMBL; AB002741; BAA21986.1; -.		
DR	InterPro; IPR000004; SppB.		
DR	SMART; SM00118; SAPB; 1.		
FT	NON_TER 1		
FT	NON_TER 87 87		
SO	SEQUENCE 87 AA; 9131 MW; FDDICIB0FA8B806E CRC64;		

Query Match	33.3%	Score 46	DB 5	Length 87
Best Local Similarity	37.0%	Pred. No. 0.71		
Matches	10	Conservative	4	Mismatches 9
				Indels 4
				Gaps 1

```
Oy      1 GFATLCTKXXXXVLDFGIDKXXXXXLI 27
          | : ||| | : ||| |
Db      59 GLVATLCTP---IASFGIAKLASIL 81
```

RESULT 4

ID	Q9PH99	PRELIMINARY:	PRC:	215 AA.
AC	Q9PH99.			
DT	01-OCT-2000	(TREMBLrel. 15.	Created)	
DT	01-OCT-2000	(TREMBLrel. 15.	Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19.	Last annotation update)	
DE	Hypothetical protein Cj0618.			
CN	Cj0618.			
OS	Campylobacter jejuni.			
OC	Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;			
CC	Campylobacter. }			
OX	NCBI_TaxID:197;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NCCTC 11168;			
RX	MEDLINE=20150912; PubMed=10688204;			
RA	Parikh1 J., Wren B.W., Mungall K., Kelsey J.M., Churcher C.,			
RA	Basham D., Chillingworth T., Davies S.R.M., Felwell T., Holroyd S.,			
RA	Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,			
RA	Quell M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,			
RA	Whitehead S., Barrett B.G.;			
RT	"The genome sequence of the food-borne pathogen Campylobacter jejuni			
RT	reveals hypervariable sequences.";			
RL	Nature 403:665-668(2000).			
DR	EMBL: AL139075; CAB75254.1; -			
KM	Hypothetical protein; Complete proteome.			
SO	SEQUENCE 215 AA; 25544 MW; 86624b0d71729d1ec CRC64;			

Query Match	31.9%	Score 44	DB 16	length 215
Best Local Similarity	40.0%	Pred. No. 4.2		
Matches	8	Conservative	4	Mismatches 8
				Indels 0
				Gaps 0
QY	1	GRIATLCTKXXXXVDFGID	20	
		:		
		:		
		:		
DB	8	GYLANLCAKIAFWSIDFILD	27	

RESULT 5			
09LS72			
ID	09LS72	PRELIMINARY;	PRT; 600 AA.
AC	09LS72		
DT	01-OCT-2000 (TREMBLrel. 15, Created)		
DT	01-OCT-2000 (TREMBLrel. 15, last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 11, last annotation update)		
DE	Gbl AAC35225.1.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Euarystota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;		
OC	Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae;		
OC	eucosids II: Brassicales; Brassicaceae; Arabidopsids.		
OX	NCBI_TaxId=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-COLUMBIA;		
RA	Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;		
RL	submitted (APR-1999) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-COLUMBIA;		
RX	MEDLINE=20277480; PubMed=10819329;		
RA	Nakamura Y.;		
RT	Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence		
RT	features of the regions of 4,504,864 bp covered by sixty p1 and TAC		
RT	clones.";		
RL	DNA Res. 7:131-135(2000).		
DR	EMBL: AB026657; BAB01819.1.		
DR	InterPro: IPR002885; PPR.		
DR	Pfam: PF01535; PPR: 11.		
DR	TIGRfams: TIGR00756; PPR: 11.		
SO	SEQUENCE 600 AA: 67338 MW: A7BD9C89BE2560D2 CRC64;		

Query Match	31.98;	Score 44;	DB 10;	Length 600;
Best Local Similarity	47.48;	Pred. NO. 12;		
Matches	9;	Conservative	3;	Mismatches 7;
				Indels 0;
				Gaps 0;

QY 2 FIATLCTKXXXXVLDGID 20
 DB 419 FIATLCTKSCNHAGLIDEGID 437

RESULT 6
 ID 084690 PRELIMINARY; PRT: 607 AA.
 AC 084690;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE Structural viral protein.
 GN H GENE.
 OS Phocine distemper virus (PDV).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
 OX NCBI_TaxID=11240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PDV/DK88;
 RX MEDLINE=92113538; PubMed=1765768;
 RA Kovamees J., Blixenkro-Moller M., Sharma B., Orvell C., Norrby E.;
 RT "The nucleotide sequence and deduced amino acid composition of the
 RT haemagglutinin and fusion proteins of the morbillivirus phocid
 RT distemper virus.";
 RL J. Gen. Virol. 72:2959-2966(1991).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=PDV/DK88;
 RA Blixenkro-Moller M.M.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL: 236979; CA85428.1; -
 DR InterPro: IPR000665; Hem-neuramndse.
 DR Pfam: PF00423; HN; 1.
 SQ SEQUENCE 607 AA; 68962 MW; BD9FB8D3AE291C5B CRC64;

Query Match
 Best Local Similarity 31.9%; Score 44; DB 12; Length 607;
 Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 3 IATLCTKXXXXVLDGIDKXXXXLIQI 30
 DB 292 LASLCTKSTILLNLGDESONSVLVI 319

RESULT 7
 ID 056986 PRELIMINARY; PRT: 607 AA.
 AC 056986;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE H protein (Fragment).
 OS Phocine distemper virus (PDV).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.
 OX NCBI_TaxID=11240;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vos H.W., Harder T.C., Kenter M., Osterhaus A.D.M.E.;
 RT "Expression of phocid distemper virus F and H genes in vaccinia
 RT virus.";
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ224707; CA12080.1; -
 DR InterPro: IPR000665; Hem-neuramndse.
 DR Pfam: PF00423; HN; 1.
 FT NON TER 607
 SQ SEQUENCE 607 AA; 68842 MW; 622D6D5A88E179E5 CRC64;

Query Match
 Best Local Similarity 31.9%; Score 44; DB 12; Length 607;
 Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

Matches 8; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
 QY 3 IATLCTKXXXXVLDGIDKXXXXLIQI 30
 DB 292 LASLCTKSTILLNLGDESONSVLVI 319

RESULT 8
 ID 08XN50 PRELIMINARY; PRT: 388 AA.
 AC 08XN50;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Capsular polysaccharide biosynthesis protein.
 GN CPE0488.
 OS Clostridium perfringens.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;
 OC Clostridiales; Clostridiaceae; Clostridium.
 OX NCBI_TaxID=1502;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=13 / TYPE A;
 RX PubMed=11792842;
 RA Shimizu T., Ohlanti K., Hirakawa H., Ohshima K., Yamashita A.,
 RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
 RT "Complete genome sequence of Clostridium perfringens, an anaerobic
 RT flesh eater.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
 DR EMBL: AP003187; BAB80194.1; -
 DR InterPro: IPR001296; Glycos_transf_1.
 DR Pfam: PF00534; Glycos_transf_1; 1.
 KW Complete proteome.
 SQ SEQUENCE 388 AA; 45089 MW; 371FC2A13775F55A CRC64;

Query Match
 Best Local Similarity 33.3%; Score 43.5; DB 16; Length 388;
 Matches 11; Conservative 5; Mismatches 6; Indels 11; Gaps 1;

QY 1 GFATLCTKXXXXVLDGIDKXXXXLIQIEDK 33
 DB 338 GFIA-----EPDIDSIYKYLKVIDDR 359

RESULT 9
 ID 0985D7 PRELIMINARY; PRT: 807 AA.
 AC 0985D7;
 DT 01-OCT-2001 (TREMblrel. 18, Created)
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Hypothetical protein ml17724.
 GN ML17724.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP003012; BAB54125.1; -
 DR InterPro: IPR000015; Fimb_usher.
 DR Pfam: PF00577; Usher; 1.
 KW Hypothetical protein; Complete proteome.

SO SEQUENCE 807 AA; 85409 MW; 912CDA09FCD8B56 CRC64;

Query Match 31.5%; Score 43.5; DB 16; Length 807;
Best Local Similarity 42.9%; Pred. No. 20;
Matches 9; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

OY 14 VLDGIDKXX-XXLIQLEDK 33
||||| 1 1 1 1 1
DB 717 VLDGVSXKADALVSLVDGK 737

RESULT 10

O4142
ID O4142 PRELIMINARY; PRT: 162 AA.
AC O4142;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Hypothetical 18.9 kDa protein.
CN C44B12.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA "Tin-Woliam A.";
RT "The sequence of C. elegans cosmid C44B12.";
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF036692; AAB88326.2; -
KW Hypothetical protein.
SQ SEQUENCE 162 AA; 18933 MW; 5D6EB5A7E970B08 CRC64;
Query Match 31.2%; Score 43; DB 5; Length 162;
Best Local Similarity 45.0%; Pred. No. 4.9;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
OY 14 VLDGIDKXX-XXLIQLEDK 33
||||| 1 1 1 1 1
DB 61 MIFGIDKXKAMISGIEEK 80
RESULT 11
O41091
ID O41091 PRELIMINARY; PRT: 389 AA.
AC O41091;
DT 01-JAN-1998 (TREMblrel. 05, Created)
DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE PBCV-1 UDP-glucose dehydrogenase.
CN A609L.
OS Paramyxium bursaria chlorella virus 1 (PBCV-1).
OC Viruses; dsDNA viruses, no RNA stage; Phycodnaviridae; Chlorovirus.
OX NCBI_TaxID=10506;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98022962; PubMed=9356347;
RA Li Y., Lu Z., Sun L., Kopp S., Kutish G.F., Rock D.L., Van Etten J.L.;

RT "Analysis of 74 kb of DNA located at the right end of the 330-kb
RT chlorella virus PBCV-1 genome.";
RL Virology 237:360-377(1997).
RN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE=20013326; PubMed=10544099;
RA Kaiser A., Volmert M., Tholl D., Graves M.V., Gurnon J.R., Xing W.,
RA Liase A.D., Nickerson K.W., Van Etten J.L.;
RT "Chlorella virus PBCV-1 encodes a functional homospemidine
RT synthase.";
RL Virology 263:254-262(1999).
RN [3]

RP SEQUENCE FROM N.A.
RX MEDLINE=20478054; PubMed=11021991;
RA Sun L., Gurnon J.R., Adams B.J., Graves M.V., Van Etten J.L.;
RT "Characterization of a beta-1,3-glucanase encoded by chlorella virus
RT PBCV-1.";
RL Virology 276:27-36(2000).
RN [4]

RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RN [5]

RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [6]

RP SEQUENCE FROM N.A.
RA Van Etten J.L.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [7]

RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [8]

RP SEQUENCE FROM N.A.
RA Graves M.V., Van Etten J.L.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [9]

RP SEQUENCE FROM N.A.
RA Gurnon J.R., Graves M.V., Van Etten J.L.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U42580; AAC96945.1; -
DR InterPro; IPR001732; UDPG_MGDP_dh.
DR Pfam; PF00984; UDPG_MGDP_dh; 1.
DR Pfam; PF03720; UDPG_MGDP_dh; 1.
DR Pfam; PF03721; UDPG_MGDP_dh; 1.
SO SEQUENCE 389 AA; 43458 MW; ED5B5A98975426E0 CRC64;

Query Match 31.2%; Score 43; DB 12; Length 389;
Best Local Similarity 31.6%; Pred. No. 12;
Matches 12; Conservative 6; Mismatches 10; Indels 10; Gaps 2;

OY 1 GFIAFLCT-----KXXXVLDGIDKXXXXXXLIQLEDK 33
|:|:| 1 1 1 1 1 :|:|:| 1
DB 10 GVGTFACAVLLAOKNEVIYVLDSER-----VOLIKNK 42

RESULT 12

O9DBR9
ID O9DBR9 PRELIMINARY; PRT: 431 AA.
AC O9DBR9;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Syntaxin binding protein 1.
CN STXB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=C57BL/6J; TISSUE=LUNG;
RA MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinaawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arkawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Irawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Straubi F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001)
DR EMBL: AK004780; BAB23558.1; -
DR MGD: MGI:107363; Stxbp1.
DR InterPro: IPR001619; Sec1-like.
DR Pfam: PF00995; Sec1.1.
SQ SEQUENCE 431 AA; 49141 MW; 5335EF4F2FCB7F35 CRC64;

Query Match 31.2%; Score 43; DB 11; Length 431;
Best Local Similarity 45.5%; Pred. No. 13;
Matches 15; Conservative 1; Mismatches 15; Indels 2; Gaps 1;

OY 3 IATLCT--KXXXXVLDGIDKXXXXLIQIIEDK 33
ID 088290 PRELIMINARY; PRT; 594 AA.
AC 088290;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Munc18-1.
NM Munc18-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SV.
RA MEDLINE=98370997; PubMed=9705297;
RA Gotch K., Yokota H., Kikuya E., Watanabe T., Oishi M.;
RT *Genomic structure of MUNC18-1 protein, which is involved in docking
RT and fusion of synaptic vesicles in brain.*;
RL J. Biol. Chem. 273:21642-21647(1998).
DR EMBL: AB012697; BAA32486.1; -
DR EMBL: AB010151; BAA32486.1; JOINED.
DR EMBL: AB012582; BAA32486.1; JOINED.
DR EMBL: AB012583; BAA32486.1; JOINED.
DR EMBL: AB012584; BAA32486.1; JOINED.
DR EMBL: AB012585; BAA32486.1; JOINED.
DR EMBL: AB012586; BAA32486.1; JOINED.
DR EMBL: AB012587; BAA32486.1; JOINED.
DR EMBL: AB012588; BAA32486.1; JOINED.
DR MGD: MGI:107363; Stxbp1.
DR MGD: MGI:2137748; Msi0g.
DR InterPro: IPR001619; Sec1-like.
DR Pfam: PF00995; Sec1.1.
SQ SEQUENCE 594 AA; 67549 MW; 5FD851D8E380079E CRC64;

```

```

Query Match 31.2%; Score 43; DB 11; Length 594;
Best Local Similarity 45.5%; Pred. No. 18;
Matches 15; Conservative 1; Mismatches 15; Indels 2; Gaps 1;

OY 3 IATLCT--KXXXXVLDGIDKXXXXLIQIIEDK 33
ID 088290 PRELIMINARY; PRT; 594 AA.
AC 088290;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Synaptaxin binding protein 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DKA/2J; TISSUE=BRAIN;
RA Fehr C., Buck K.J.;
RT *Characterization of candidate genes for multiple ethanol influenced
RT traits on mouse chromosome 2.*;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF326545; AAL37391.1; -
DR InterPro: IPR001619; Sec1-like.
DR Pfam: PF00995; Sec1.1.
SQ SEQUENCE 594 AA; 67567 MW; 281F1070E3BCE251 CRC64;

Query Match 31.2%; Score 43; DB 11; Length 594;
Best Local Similarity 45.5%; Pred. No. 18;
Matches 15; Conservative 1; Mismatches 15; Indels 2; Gaps 1;

OY 3 IATLCT--KXXXXVLDGIDKXXXXLIQIIEDK 33
ID 176 IATLCATLKEYPAYRYRGEXKDNALLAQIIODK 208
DB 176 IATLCATLKEYPAYRYRGEXKDNALLAQIIODK 208

RESULT 15
OY 096TC8 PRELIMINARY; PRT; 603 AA.
ID 096TC8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE HUNC18p.
NM Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98207254; PubMed=9545644;
RA Swanson D.A., Steel J.M., Vallie D.;
RT STXP1, a protein implicated in vesicle trafficking and
RT neurotransmitter release.*;
RL Genomics 48:373-376(1998).
DR EMBL: AF004563; AAC39689.1; -
DR InterPro: IPR001619; Sec1-like.
DR Pfam: PF00995; Sec1.1.
SQ SEQUENCE 603 AA; 68735 MW; 73D8337768259493 CRC64;

Query Match 31.2%; Score 43; DB 11; Length 603;
Best Local Similarity 45.5%; Pred. No. 18;
Matches 15; Conservative 1; Mismatches 15; Indels 2; Gaps 1;

OY 3 IATLCT--KXXXXVLDGIDKXXXXLIQIIEDK 33
ID 096TC8 PRELIMINARY; PRT; 603 AA.
AC 096TC8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE HUNC18p.
NM Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98207254; PubMed=9545644;
RA Swanson D.A., Steel J.M., Vallie D.;
RT STXP1, a protein implicated in vesicle trafficking and
RT neurotransmitter release.*;
RL Genomics 48:373-376(1998).
DR EMBL: AF004563; AAC39689.1; -
DR InterPro: IPR001619; Sec1-like.
DR Pfam: PF00995; Sec1.1.
SQ SEQUENCE 603 AA; 68735 MW; 73D8337768259493 CRC64;

```

Mon May 12 15:29:49 2003

us-09-851-422a-1.rpt

Page 6

Db 176 IATLCATLKEYPAVRNGEYKDNALLAQLIQDK 208

Search completed: May 9, 2003, 15:17:34
Job time : 29.3922 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:07:45 ; Search time 53.6471 Seconds
(without alignments)
94.386 Million cell updates/sec

Title: US-09-851-422A-2

Perfect score: 141

Sequence: 1 GIGAVLKKXXXVLTGTPALISWIKXXXXRXXXXRQ 38

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 13250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A.Geneseq.101002.*
1: /SID2/gcgdata/geneseq/genesep-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/genesep-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/genesep-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/genesep-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/genesep-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/genesep-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/genesep-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/genesep-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/genesep-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/genesep-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/genesep-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/genesep-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/genesep-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/genesep-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/genesep-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/genesep-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/genesep-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/genesep-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/genesep-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/genesep-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/genesep-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/genesep-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/genesep-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	66.7	28	23	AAE18202
2	94	66.7	36	14	AAE18202
3	94	66.7	36	17	AAW01430
4	94	66.7	36	20	AAV49867
5	94	66.7	36	21	AAE13927
6	94	66.7	36	21	AAV77660
7	94	66.7	36	22	AAU11168
8	93	66.0	26	10	AAE1340
9	93	66.0	26	12	AAE13908
10	93	66.0	26	14	AAE45114

11	93	66.0	26	14	AAE35383	Melittin used to t
12	93	66.0	26	14	AAE39759	Melittin peptide.
13	93	66.0	26	15	AAE55989	Ion channel form
14	93	66.0	26	15	AAE50565	Amphiphilic pep
15	93	66.0	26	15	AAE50430	Amphiphilic pep
16	93	66.0	26	15	AAE56950	Peptide which neu
17	93	66.0	26	15	AAE59067	Melittin, cancer t
18	93	66.0	26	16	AAE90136	Melittin modified
19	93	66.0	26	16	AAE85516	Melittin. Synthet
20	93	66.0	26	16	AAE72873	Calmodulin-binding
21	93	66.0	26	17	AAE08667	Honeybee venom mel
22	93	66.0	26	18	AAE35145	Melittin-derived p
23	93	66.0	26	18	AAE35146	Melittin-derived p
24	93	66.0	26	18	AAE09134	Melittin antihac
25	93	66.0	26	18	AAE16374	Honeybee peptide m
26	93	66.0	26	19	AAE82879	Antipathogenic pep
27	93	66.0	26	19	AAE82880	Antipathogenic pep
28	93	66.0	26	19	AAE66453	Cationic peptide m
29	93	66.0	26	19	AAE71674	Melittin-derived t
30	93	66.0	26	19	AAE77385	Lytic peptide with
31	93	66.0	26	19	AAE43128	Melittin, the acti
32	93	66.0	26	20	AAE22019	Melittin. Homo sa
33	93	66.0	26	20	AAE10732	Peptide used to ma
34	93	66.0	26	20	AAE95333	Synthetic bee mel
35	93	66.0	26	21	AAE11034	Apis mellifera mel
36	93	66.0	26	21	AAE17407	Antipathogenic pep
37	93	66.0	26	21	AAE17408	Antipathogenic pep
38	93	66.0	26	21	AAE17409	Antipathogenic pep
39	93	66.0	26	21	AAE12439	Plasmaid construct
40	93	66.0	26	21	AAE91752	Cationic peptide M
41	93	66.0	26	21	AAE44325	Antimicrobial pep
42	93	66.0	26	22	AAE92169	Signal transductio
43	93	66.0	26	22	AAE72458	Melittin to determ
44	93	66.0	26	22	AAE50842	Bee protein calmod
45	93	66.0	26	22	AAE50199	Membrane disruptiv

ALIGNMENTS

RESULT 1	
AAE18202	
ID	AAE18202 standard; peptide: 28 AA.
XX	
AC	AAE18202:
XX	
DT	07-MAY-2002 (first entry)
DE	Procytotoxin #4 for treating cancer.
XX	
KW	Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
KW	prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX	
OS	Unidentified.
XX	
FH	Key
FT	Modified-site
FT	Location/Qualifiers
FT	21..23
FT	/note- "Lys([epsilon-gamma]-Glu)-Arg; where
FT	[epsilon-gamma] represents a peptide bond between the
FT	epsilon amino group of lysine and the gamma carboxyl
FT	group of adjacent glutamate, and a standard peptide
FT	Linkage between lysine and arginine"
FT	24..26
FT	/note- "Lys([epsilon-gamma]-Glu)-Arg; where
FT	[epsilon-gamma] represents a peptide bond between the
FT	epsilon amino group of lysine and the gamma carboxyl
FT	group of adjacent glutamate, and a standard peptide
FT	Linkage between lysine and arginine"
PN	WO200185777-A2.
XX	
PD	15-NOV-2001.
XX	

PF 09-MAY-2001; 2001WO-US40690.
 XX 09-MAY-2000; 2000US-203063P.
 PR 16-JUN-2000; 2000US-212042P.
 XX
 PA (GREE-) GREENVILLE HOSPITAL SYSTEM.
 XX
 PI Yu X, Wagner TE;
 XX
 DR WPI: 2002-154423/20.
 XX
 PT A procytotoxin comprising a cytotoxic peptide with at least one lysine
 PT residue which is converted into a cytotoxin in a target cell is used
 PT for treating cancer, particularly of the prostate, skin, ovary or lung
 PT
 PS Disclosure: Page 12; 33pp; English.
 XX
 CC The present invention relates to a procytotoxin, comprising a cytotoxic
 CC peptide with at least one lysine residue bound by a peptide bond to at
 CC least one amino acid via the epsilon-amino group of the lysine residue.
 CC The procytotoxin is used to treat cancer, particularly of the prostate,
 CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
 CC peptide, typically disrupt cell membrane, causing cell lysis and death
 CC (apoptosis) upon contact. The present sequence is procytotoxin.
 CC
 SQ Sequence 28 AA;
 XX
 Query Match 66.7%; Score 94; DB 23; Length 28;
 Best Local Similarity 71.0%; Pred. No. 9.1e-09;
 Matches 22; Conservative 1; Mismatches 4; Indels 4; Gaps 1;
 OY 1 GIGAVLKXXXVLTGTPALISWIKKXXXRK 31
 ||||| |||||
 DB 1 GIGAVLK---VLTGTPALISWIKRKRRQ 27
 ||||| |||||
 RESULT 2
 AAR34603
 ID AAR34603 standard; peptide; 36 AA.
 XX
 AC AAR34603;
 XX
 DT 09-AUG-1993 (first entry)
 XX
 DE Sequence of endosomolytic peptide.
 XX
 DE Endosomolytic peptide; endosome; transfection; gene transfer.
 XX
 OS Synthetic.
 XX
 KM WO9307283-A.
 XX
 PN 15-APR-1993.
 XX
 PD 28-SEP-1992; 92WO-EP02234.
 XX
 PF 30-SEP-1991; 91US-0767788.
 XX
 PR 30-SEP-1991; 91US-0768039.
 PR 30-JAN-1992; 92US-0827102.
 PR 30-JAN-1992; 92US-0827103.
 PR 07-APR-1992; 92US-0864759.
 PR 02-SEP-1992; 92US-0937788.
 XX
 PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 PA (GETH) GENENTECH INC.
 PA (UNIC-) UNIV NORTH CAROLINA.
 XX
 PI Birstiel ML, Cotten M, Curiel D, Oberhauser B;
 PI Plank C, Schmidt MCM, Wagner E, Zatloukal K;
 XX
 DR WPI: 1993-134470/16.
 XX

PT Compn. for efficient transfection of higher eukaryotic cells -
 PT contains conjugate of nucleic acid, and affinity substance plus
 PT endosomolytic agent, esp. adenovirus, useful e.g. in gene
 PT transfer therapy
 XX
 PS Example: Page 196; 281pp; German.
 XX
 CC The peptide is endosomolytic. It can release the contents of
 CC endosomes into the cytoplasm. It can be covalently or non-
 CC covalently bound to a nucleic acid affinity substance (NAS) directly
 CC or via an NA-binding domain (NA - nucleic acid). The compns. are
 CC esp. used to introduce NA into human cells.
 CC
 SQ Sequence 36 AA;
 XX
 Query Match 66.7%; Score 94; DB 14; Length 36;
 Best Local Similarity 71.0%; Pred. No. 1.2e-08;
 Matches 22; Conservative 1; Mismatches 4; Indels 4; Gaps 1;
 OY 1 GIGAVLKXXXVLTGTPALISWIKKXXXRK 31
 ||||| |||||
 DB 1 GIGAVLK---VLTGTPALISWIKRKRRQ 27
 ||||| |||||
 RESULT 3
 AAW01430
 ID AAW01430 standard; peptide; 36 AA.
 XX
 AC AAW01430;
 XX
 DT 27-JAN-1997 (first entry)
 XX
 DE Melittin peptide 1 with C-terminal oligo-Lys tail.
 XX
 DE Transfection: gene therapy; non-viral peptide; internalising factor;
 KM endosome; cytoplasm; oligolysin.
 XX
 OS Synthetic.
 XX
 PN US547932-A.
 XX
 PD 20-AUG-1996.
 XX
 PF 30-SEP-1991; 91US-0767788.
 XX
 PR 23-SEP-1992; 92US-0948357.
 PR 30-SEP-1991; 91US-0767788.
 PR 30-SEP-1991; 91US-0768039.
 PR 30-JAN-1992; 92US-0827102.
 PR 30-JAN-1992; 92US-0827103.
 PR 07-APR-1992; 92US-0864759.
 PR 02-SEP-1992; 92US-0937788.
 XX
 PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 PA (GETH) GENENTECH INC.
 XX
 PI Birstiel ML, Cotten M, Curiel DT, Oberhauser B;
 PI Plank C, Schmidt MCM, Wagner E, Zatloukal K;
 XX
 DR WPI: 1996-392650/39.
 XX
 PT Introduction of nucleic acid into higher eukaryotic cells - as
 PT complex with nucleic acid affinity agent and endosomolytic agent
 PT
 PS Example 37; Column 95-96; 109pp; English.
 XX
 CC A nucleic acid can be introduced into the cytoplasm of a higher
 CC eukaryotic cell by contacting the cell with a compn. comprising a
 CC complex, where the complex comprises a nucleic acid of interest, a
 CC substance with affinity for the nucleic acid and an endosomolytic
 CC agent, which allows internalisation of the complex into the cell by
 CC endocytosis, and is released into the cytoplasm by endosomolysis.
 CC The present sequence is a melittin peptide designated mel 1, which

CC is used in the transfection of hepatocytes.
 XX
 SQ Sequence 36 AA;
 Query Match 66.7%; Score 94; DB 17; Length 36;
 Best Local Similarity 71.0%; Pred. No. 1.2e-08;
 Matches 22: Conservative 1; Mismatches 4; Indels 4; Gaps 1;
 Oy 1 GIGAVLKXXXVLTGTPALISMIRKXXXRK 31
 ||||| |||||
 Db 1 GIGAVLK-----VLTGTPALISMIRKRRQK 27

RESULT 4
 AAY49867
 ID AAY49867 standard; peptide: 36 AA.
 AC AAY49867;
 XX
 DT 21-JAN-2000 (first entry)
 XX
 DE Melittin peptide mel 1 with an oligolysine C-terminal extension.
 XX
 KW Endosomolytic; influenza; haemagglutinin; gene transfer; transferrin;
 transfection; gene therapy.
 XX
 OS Synthetic.
 XX
 PN US5981273-A.
 PD 09-NOV-1999.
 XX
 PF 25-MAY-1995; 95US-0450417.
 XX
 PR 23-SEP-1992; 92US-0948357.
 PR 30-SEP-1991; 91US-076788.
 PR 30-SEP-1991; 91US-0768039.
 PR 30-JAN-1992; 92US-0827102.
 PR 30-JAN-1992; 92US-0827103.
 PR 07-APR-1992; 92US-0864759.
 PR 02-SEP-1992; 92US-0937788.
 XX
 PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 PA (GETH) GENENTECH INC.
 PA (UYNC-) UNIV NORTH CAROLINA.
 XX
 PI Zatloukal K, Wagner E, Schmidt WGM, Birnstiel ML, Curiel DT;
 PI Plank C, Oberhauser B, Cotten M;
 XX
 DR WPI: 1999-633329/54.
 XX
 PT Composition for transfection of eukaryotic cells, useful in gene
 PT therapy -
 XX
 PS Example 37: Column 76-77; 117pp: English.
 XX
 CC The present invention describes a composition for the transfection of a
 CC higher eukaryotic cell which comprises a nucleic acid of interest and an
 CC endosomolytic agent. The composition for the transfection of a higher
 CC eukaryotic cell comprises a non-naturally occurring nucleic acid
 CC complex, where the nucleic acid complex comprises: (a) a nucleic acid of
 CC interest; (b) a substance which binds to nucleic acid; and (c) an
 CC endosomolytic agent which is bound to the substance of (b), where the
 CC endosomolytic agent is a virus or a virus component which is not an
 CC internalising factor for a human cell and where the complex is
 CC internalised into the cell by endocytosis, and is released into the
 CC cell by endosomolysis. The composition is useful for delivering specific
 CC genes into a cell and therefore has applications in gene therapy. Since
 CC the method of gene transfer is a physiological method, represented by
 CC receptor-mediated endocytosis, it is non-toxic compared to the prior art
 CC (viral vectors). The present sequence represents a melittin peptide
 CC with an oligolysine C-terminal extension, which is used in an
 CC example from the present invention.

XX
 SQ Sequence 36 AA;
 Query Match 66.7%; Score 94; DB 20; Length 36;
 Best Local Similarity 71.0%; Pred. No. 1.2e-08;
 Matches 22: Conservative 1; Mismatches 4; Indels 4; Gaps 1;
 Oy 1 GIGAVLKXXXVLTGTPALISMIRKXXXRK 31
 ||||| |||||
 Db 1 GIGAVLK-----VLTGTPALISMIRKRRQK 27

RESULT 5
 AAB13927
 ID AAB13927 standard; peptide: 36 AA.
 AC AAB13927;
 XX
 DT 15-NOV-2000 (first entry)
 XX
 DE Melittin peptide mel 1.
 XX
 KW Endosomolytic; gene therapy; transfection; melittin; mel 1.
 XX
 OS Unidentified.
 XX
 PN US6077663-A.
 PD 20-JUN-2000.
 XX
 PF 25-MAY-1995; 95US-0449754.
 XX
 PR 23-SEP-1992; 92US-0948357.
 PR 30-SEP-1991; 91US-076788.
 PR 30-SEP-1991; 91US-0768039.
 PR 30-JAN-1992; 92US-0827102.
 PR 30-JAN-1992; 92US-0827103.
 PR 07-APR-1992; 92US-0864759.
 PR 02-SEP-1992; 92US-0937788.
 XX
 PA (BOEH) BOEHRINGER INGELHEIM INT GMBH.
 PA (GETH) GENENTECH INC.
 PA (UYNC-) UNIV NORTH CAROLINA.
 XX
 PI Zatloukal K, Wagner E, Schmidt WGM, Birnstiel ML, Curiel DT;
 PI Plank C, Oberhauser B, Cotten M;
 XX
 DR WPI: 2000-450936/39.
 XX
 PT Novel composition for introducing nucleic acid complexes into higher
 PT eukaryotic cells comprises a non-naturally occurring nucleic acid
 PT complex containing the nucleic acid of interest and an endosomolytic
 PT agent -
 XX
 PS Example 37: Column 74; 114pp: English.
 XX
 CC The present invention relates to a composition for the transfection of
 CC higher eukaryotic cells. The nucleic acid of interest is contained in an
 CC endosome which is internalised into the cells. The nucleic acid is then
 CC released into the cytoplasm from the endosome by an endosomolytic agent.
 CC The present sequence is the melittin peptide with an oligo lysine
 CC tail, designated mel 1. The composition of the invention is useful in
 CC assays for determining the host immune response to a given antigen and
 CC also in gene therapy. The composition is broadly applicable making it
 CC possible to make use of a positive effect of viruses to cells which do
 CC not have any receptor for the virus.
 XX
 SQ Sequence 36 AA;
 Query Match 66.7%; Score 94; DB 21; Length 36;
 Best Local Similarity 71.0%; Pred. No. 1.2e-08;
 Matches 22: Conservative 1; Mismatches 4; Indels 4; Gaps 1;

```

QY 1 GIGAVLKXXXXVLTGTPALISWIKXXXXXRR 31
      ||||| ||||| : |
DB 1 GIGAVLK----VLTGTPALISWIKRRRQOK 27

```

RESULT 6
AA777660

Query Match	66.7%	Score 94:	DB 21:	Length 36:
Best Local Similarity	71.0%	Pred. No.	1.2e-08:	
Matches 22:	Conservative	1:	Mismatches 4:	Indels 4:
				Gaps 1:

```

Qy      1 GIGAVLKXXXXVLTGTPALISWIKXXXXRK 31
        |||||  |||||
Db      1 GIGAVLK----VLTGTPALISWIKRRRQK 27

```

```

RESULT 7
AAU11168
ID   AAU11168 standard; peptide; 36 AA

```

Query Match 66.78; Score 94; DB 22; Length 36;

OS Synthetic.

PN W09324138-A.
 XX
 XX 09-DEC-1993.
 XX
 XX 27-MAY-1993: 93WO-US05192.
 PF
 XX 01-JUN-1992: 92US-0891201.
 PR
 XX (MAGA-) MAGAININ PHARM INC.
 PA
 XX Karl U;
 PI
 XX WPI: 1993-405419/50.
 DR
 XX Peptide(s) or proteins with an N-terminal lipophilic substt.
 PT used for inhibiting growth of target cell, virus or
 PT virally-infected cell
 PS
 XX Disclosure: Page 39, 113pp: English.
 PS
 XX A novel compsn. for inhibiting growth of a target cell, virus or
 CC virally-infected cell comprises a peptide of formula T-N(W)-X (I).
 CC X is a biologically active amphiphilic ion channel-forming peptide
 CC or protein; pref. a magainin peptide, a PGla peptide, a XPe
 CC peptide, a CPF peptide, a cecropin or a sarcotoxin.
 CC N is the nitrogen of the N-terminal amino group.
 CC T is a lipophilic moiety; pref. R-CO, where R is a 2-16C
 CC hydrocarbon (alkyl or aromatic or alkylaromatic).
 CC T is pref. an octanoyl group.
 CC W is T or hydrogen.
 CC Melittin is an amphiphatic peptide and an example of X.
 CC The N-terminal substt. peptides and proteins have increased
 CC biological activity as compared with unsubst. peptides or proteins
 CC or peptides substt. at the N-terminal with an acetyl gp.
 CC They can be used as antimicrobial agents, antiviral agents,
 CC antitumor agents, antiparasitic agents or spermicides and
 CC can also exhibit other bioactive functions. They can also be
 CC used in promoting or stimulating wound healing, for the treatment
 CC of external burns and to treat and/or prevent skin and burn
 CC infections or eye infections.
 CC
 CC
 SO Sequence 26 AA:
 Query Match 66.0%; Score 93; DB 14; Length 26;
 Best Local Similarity 68.4%; Pred. No. 1.2e-08;
 Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;
 Oy 1 GIGAVLKXXXVLTGTPALISWIKXXXXRXXRQ 38
 Db 1 GIGAVLK---VLTGTPALISWIK---RK---RQ 26
 RESULT 11
 AAR35383
 ID AAR35383 standard; peptide: 26 AA.
 XX
 XX AAR35383;
 AC
 XX
 XX 07-JUN-1993 (first entry)
 DT
 XX
 XX Melittin used to treat oral infections.
 DE
 XX Adverse oral conditions; amphipathic; anti-bacterial; anti-viral;
 KW anti-fungal; dental plaque; dental caries; periodontal disease;
 KW gingivitis; ionophore; ion-channel forming; honeybee; venom.
 XX
 OS Apis mellifera.
 XX
 XX W09301723-A.
 PN
 XX
 XX 04-FEB-1993.
 PD
 XX
 XX 09-JUL-1992: 92WO-US05757.
 PF

XX
 XX 25-JUL-1991: 91US-0735070.
 PR
 XX
 XX (MAGA-) MAGAININ PHARM INC.
 PA
 XX
 XX Berkowitz B, Jacob L;
 PI
 XX WPI: 1993-058434/07.
 DR
 XX
 XX Peptide(s) for prophylaxis and treatment of oral disorders - used
 PT for periodontal disease, plaque, dental caries, gingivitis, etc.
 PT
 XX
 PS Claim 2: Page 130; 143pp: English.
 PS
 XX Melittin is an amphipathic peptide which is isolated from honeybee
 CC venom. The peptide is known to be cytolytic (see Hoppe-Seyster's
 CC Zeitschrift Physiol. Chem. 348, 37-50, 1987) and is a preferred
 CC peptide for use in preventing or treating adverse oral conditions.
 CC The peptide is suitable for use in oral compositions to treat or
 CC prevent periodontal disease, plaque, dental caries, halitosis and
 CC gingivitis. The anti-bacterial action will also be useful against
 CC bacteria associated with dental implant infections and the peptides
 CC can stimulate the healing of wounds in the oral cavity.
 CC
 CC
 SO Sequence 26 AA:
 Query Match 66.0%; Score 93; DB 14; Length 26;
 Best Local Similarity 68.4%; Pred. No. 1.2e-08;
 Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;
 Oy 1 GIGAVLKXXXVLTGTPALISWIKXXXXRXXRQ 38
 Db 1 GIGAVLK---VLTGTPALISWIK---RK---RQ 26
 RESULT 12
 AAR39759
 ID AAR39759 standard; peptide: 26 AA.
 XX
 XX AAR39759;
 AC
 XX
 XX 08-FEB-1994 (first entry)
 DT
 XX
 XX Melittin peptide.
 DE
 XX Melittin: amphipathic; honeybee venom; cytolytic; antifungal;
 KW antibiotic; antiparasitic; antimicrobial; spermicidal; antiviral;
 KW
 XX Apis mellifera.
 OS
 XX US5235038-A.
 PN
 XX
 XX 10-AUG-1993.
 PD
 XX
 XX 22-JAN-1991: 91US-0643343.
 PF
 XX
 XX 22-JAN-1991: 91US-0643343.
 PR
 XX (TORR-) TORRY PINES INST MOLECULAR STUDIES.
 PA
 XX
 XX Blondelle SE, Houghten RA;
 PI
 XX
 XX WPI: 1993-264689/33.
 DR
 XX
 XX Deletion and substitution peptide analogues of melittin - used as
 PT antibiotics, antifungal agents, antitumor agents, antiviral
 PT agents or to stimulate wound healing
 PT
 XX
 PS Disclosure: Page 1; 36pp: English.
 PS
 XX The peptide melittin (AAR39759) from honeybee venom was used to
 CC produce deletion analogues (AAR39759-R39783) and substitution
 CC analogues (AAR39784-R39890) (AAR38827-R38839). The peptides can be used

CC to inhibit growth of the target cell, thus exhibiting antimicrobial,
 CC anti-viral, antibiotic, anti-tumor, antiparasitic, antifungal and
 CC spermicidal properties.

XX Sequence 26 AA:

Query Match 66.0%; Score 93; DB 14; Length 26;
 Best Local Similarity 68.4%; Pred. No. 1.2e-08;
 Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

OY 1 GIGAVLXXXXVLTGTPALISWIKXXXXRQ 38
 DB 1 GIGAVLK-----VLTGTPALISWIK----RK----RQ 26

RESULT 13

AAR55989 ID AAR55989 standard; peptide: 26 AA.

XX AAR55989;

DT 19-DEC-1994 (first entry)

DE Ion channel forming peptide (mellitin).

KW Ion channel forming peptide; tumour; skin disease; malignancy;
 KW melanoma; carcinoma; basal cell; squamous cell; magainin; pGla;
 KW CYP peptides; cecropins; sarcotoxin; mellitin; apidocin; defensins;
 KW major basic protein; bacteria-permeability increasing protein;
 KW perforin.

XX Apis mellifera.

XX WO9412206-A.

XX 09-JUN-1994.

XX 03-DEC-1993; 93WO-US11885.

XX 03-DEC-1992; 92US-0984957.

PA (MAGA-) MAGAININ PHARM INC.

PI Jacob LS, Maloy WL;

WPI: 1994-199965/24.

PT Treating skin cancer with ion channel forming peptide(s) - e.g.
 PT magainins, mellitin etc., specifically for treating melanoma
 PS Disclosure: Page 122; 136pp; English.

CC The peptide is used to treat dermatological malignancies. It
 CC may be used to treat especially melanoma but also basal cell and
 CC squamous cell carcinomas. It can be used together with an ion which
 CC also inhibits/prevents growth of the target cell. Peptides used for
 CC such therapy include magainin, pGla or CYP peptides; cecropins,
 CC sarcotoxins, mellitin, apidocins, defensins, major basic protein of
 CC eosinophils; bacteria-permeability increasing protein and perforin.
 CC See also AAO55876-Q55997.

XX Sequence 26 AA:

Query Match 66.0%; Score 93; DB 15; Length 26;
 Best Local Similarity 68.4%; Pred. No. 1.2e-08;
 Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

OY 1 GIGAVLXXXXVLTGTPALISWIKXXXXRQ 38
 DB 1 GIGAVLK-----VLTGTPALISWIK----RK----RQ 26

RESULT 14

AAR50565 ID AAR50565 standard; peptide: 26 AA.

XX AAR50565;

DT 18-OCT-1994 (first entry)

DE Amphiphilic peptide #114.

KW Amphiphilic; ion forming; gynaecological malignancy; magainin; pGla;
 KW XPF; CYP; cecropin; sarcotoxin; mellitin; apidocin; defensin;
 KW major basic protein; eosinophils; uterine; cervical; cancer;
 KW bacterial permeability increasing protein; ovarian; stage IC.

OS Synthetic.

PN WO9405313-A.

PD 17-MAR-1994.

PF 16-AUG-1993; 93WO-US07798.

PR 31-AUG-1992; 92US-0937462.

PA (MAGA-) MAGAININ PHARM INC.

PI Baker MA, Jacob LS, Maloy WL;

WPI: 1994-100851/12.

PT Treating gynaecological tumours with amphiphilic peptide(s) -
 PT which form ion channels, e.g. magainin or pGla peptide(s),
 PT partic. for treating ovarian, uterine or cervical cancers.
 PS Disclosure: Page 117; 130pp; English.

CC The sequences given in AAR50452-568 represent amphiphilic, ion forming
 CC peptides which may be used to treat gynaecological malignancy.
 CC These peptides are based on magainin, pGla, XPF or CYP, a cecropin, a
 CC sarcotoxin, mellitin, an apidocin, a defensin, major basic protein of
 CC eosinophils or a bacterial permeability increasing protein. These
 CC peptides are esp. used to treat ovarian, esp. stage IC, uterine or
 CC cervical cancers.

XX Sequence 26 AA:

Query Match 66.0%; Score 93; DB 15; Length 26;
 Best Local Similarity 68.4%; Pred. No. 1.2e-08;
 Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

OY 1 GIGAVLXXXXVLTGTPALISWIKXXXXRQ 38
 DB 1 GIGAVLK-----VLTGTPALISWIK----RK----RQ 26

RESULT 15

AAR50430 ID AAR50430 standard; peptide: 26 AA.

XX AAR50430;

DT 17-OCT-1994 (first entry)

DE Amphiphilic peptide #95, mellitin derivative.

KW Amphiphilic peptide; aprotic organic solvent; alcohol; antitumour;
 KW antibiotic; antimicrobial; antifungal; antiparasitic; anticancer;
 KW antiviral; human; animal; plant; ion-channel; forming peptide.

OS Apis mellifera.

PN WO9405308-A.

XX

PD 17-MAR-1994

PF 13-AUG-1993; 93WO-US07694.

PR 28-AUG-1992; 92US-0936504.

PA (MAGN-) MAGNIN PHARM INC.
xx

PI Williams JI;

DR WPI; 1994-100846/12.

PT Purifying amphiphilic protein or peptide by solvent extn. -
PT partic. for recombinant, ion-channel forming peptide(s) such as
PT magainins, avoids use of chaotropic agents.

PS Disclosure; Page 117; 135pp; English.

CC The sequences given in AAR50336-451 are amphiphilic peptides which
CC were isolated by the method of the invention. A material containing
CC amphiphilic peptides such as these, was treated with a mixt. of
CC aprotic organic solvent and alcohol to form a single miscible
CC solution. This solution was then treated with a aqueous solution to
CC form an aqueous phase solution containing the peptides and an
CC organic solvent, and the peptides were isolated from the
CC aqueous phase. The isolated peptides may be useful as antibiotic,
CC antimicrobial, antifungal, antiparasitic, antitumor, anticancer,
CC and/or antiviral agents for treatment of humans, animals or plants
CC These peptides are esp. ion-channel forming peptides which enable
CC biologically active ions to enter cells.

SQ Sequence 26 AA;

Query Match	66.08;	Score 93;	DB 15;	Length 26;
Post Local Similarity	59.48;	Prod No. 1	2e-08	

Matches	26;	Conservative	0;	Mismatches	0;	Indels	12;	Gaps	3;
---------	-----	--------------	----	------------	----	--------	-----	------	----

Qy 1 GIGAVLKKXXVLTGTPALISWIKKXKXXXXRQQ 38

Db 1 GIGAVLK---VLTGGLPALISWIK---RK---RQQ 26

Search completed: May 9, 2003, 16:15:35
Job time : 54.6471 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:13:10 ; Search time 13.7843 Seconds
(without alignments)
81.112 Million cell updates/sec

Title: US-09-851-422A-2

Perfect score: 141
Sequence: 1 GIGAVLXXXXXVLTGTPALISWIKXXXXRRXXXXRQ 38Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	66.7	36	1	US-07-948-357-12 Sequence 12, Appl
2	94	66.7	36	2	US-08-450-417-12 Sequence 12, Appl
3	94	66.7	36	3	US-08-449-741-12 Sequence 12, Appl
4	94	66.7	36	4	US-08-449-754-12 Sequence 12, Appl
5	94	66.7	36	4	US-09-465-646-12 Sequence 12, Appl
6	93	66.0	26	1	US-07-643-343A-1 Sequence 1, Appl
7	93	66.0	26	1	US-07-626-589-3 Sequence 3, Appl
8	93	66.0	26	1	US-08-079-445-7 Sequence 3, Appl
9	93	66.0	26	1	US-07-943-709-58 Sequence 58, Appl
10	93	66.0	26	1	US-08-236-410-3 Sequence 3, Appl
11	93	66.0	26	1	US-08-218-608-8 Sequence 8, Appl
12	93	66.0	26	1	US-08-434-120-114 Sequence 114, App
13	93	66.0	26	1	US-08-465-325-93 Sequence 93, Appl
14	93	66.0	26	1	US-08-434-761-7 Sequence 7, Appl
15	93	66.0	26	1	US-08-039-557-5 Sequence 5, Appl
16	93	66.0	26	1	US-08-519-180-11 Sequence 11, Appl
17	93	66.0	26	1	US-08-467-623-1 Sequence 1, Appl
18	93	66.0	26	1	US-08-467-623-2 Sequence 2, Appl
19	93	66.0	26	1	US-08-465-421-3 Sequence 3, Appl
20	93	66.0	26	2	US-08-818-253-34 Sequence 34, Appl
21	93	66.0	26	3	US-08-782-997A-4 Sequence 4, Appl
22	93	66.0	26	3	US-08-984-294-19 Sequence 19, Appl
23	93	66.0	26	3	US-08-641-873-3 Sequence 3, Appl
24	93	66.0	26	4	US-09-125-605-2 Sequence 2, Appl
25	93	66.0	26	4	US-09-230-180-34 Sequence 34, Appl
26	93	66.0	26	4	US-08-818-252-34 Sequence 34, Appl
27	93	66.0	26	4	US-08-749-066A-7 Sequence 7, Appl

28	93	66.0	26	4	US-09-319-730-13 Sequence 13, Appl
29	93	66.0	26	4	US-09-115-737-93 Sequence 93, Appl
30	93	66.0	26	4	US-08-842-322-28 Sequence 28, Appl
31	93	66.0	26	4	US-09-316-919-50 Sequence 50, Appl
32	93	66.0	28	3	US-08-118-055-11 Sequence 11, Appl
33	93	65.2	24	2	US-08-553-271-2 Sequence 2, Appl
34	92	65.2	25	1	US-09-285-070-2 Sequence 2, Appl
35	92	65.2	25	1	US-07-643-343A-22 Sequence 22, Appl
36	92	65.2	25	1	US-07-643-343A-23 Sequence 23, Appl
37	92	65.2	25	1	US-07-643-343A-24 Sequence 24, Appl
38	92	65.2	25	1	US-07-643-343A-25 Sequence 25, Appl
39	92	65.2	26	1	US-07-643-343A-41 Sequence 41, Appl
40	92	65.2	26	1	US-07-643-343A-42 Sequence 42, Appl
41	92	65.2	26	1	US-07-643-343A-43 Sequence 43, Appl
42	92	65.2	26	1	US-07-643-343A-44 Sequence 44, Appl
43	92	65.2	26	1	US-07-643-343A-45 Sequence 45, Appl
44	92	65.2	26	1	US-08-141-324-7 Sequence 7, Appl
45	92	65.2	26	1	US-08-295-086-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-07-948-357-12
; Sequence 12, Application US/07948357
; Patent No. 5547932
; GENERAL INFORMATION:
; APPLICANT: Curtiel, David T.
; APPLICANT: Birnstiel, Max L.
; APPLICANT: Cotten, Matthew
; APPLICANT: Wagner, Ernst
; APPLICANT: Zalloukal, Kurt
; APPLICANT: Plank, Christian
; APPLICANT: Oberhauser, Berndt
; APPLICANT: Schmidt, Walter G.M.
; TITLE OF INVENTION: Composition for Introducing Nucleic Acid
; TITLE OF INVENTION: Complexes Into Higher Eucaryotic Cells
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
; STREET: 1225 Connecticut Avenue
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/948,357
; FILING DATE: 19920923
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 0652.0940004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 466-0800
; TELEFAX: (202) 833-8716
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 36 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: both
; MOLECULE TYPE: peptide
; US-07-948-357-12
; Query Match 66.7%; Score 94; DB 1; Length 36;
; Best Local Similarity 71.0%; Pred. No. 5,9e-10;

Matches 22; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

QY 1 GIGAVLXXXXVLTGTPALISWIKXXXXRK 31
||||| ||||||| :1
DB 1 GIGAVLK----VLTGTPALISWIKKKRQK 27

RESULT 2

US-08-450-417-12
Sequence 12, Application US/08450417
Patent No. 5981273

GENERAL INFORMATION:
APPLICANT: Curjel, David T.
APPLICANT: Birnstiel, Max L.
APPLICANT: Cotten, Matthew
APPLICANT: Wagner, Ernst
APPLICANT: Zatloukal, Kurt
APPLICANT: Plank, Christian
APPLICANT: Oberhauser, Berndt
APPLICANT: Schmidt, Walter G.M.
TITLE OF INVENTION: Composition for Introducing Nucleic Acid
TITLE OF INVENTION: Complexes Into Higher Eucaryotic Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,417

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/948,357
FILING DATE: 19920923
ATTORNEY/AGENT INFORMATION:
NAME: Esmont, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0652,0940004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: both
MOLECULE TYPE: peptide
US-08-450-417-12

Query Match 66.7%; Score 94; DB 2; Length 36;
Best Local Similarity 71.0%; Pred. No. 5.9e-10;
Matches 22; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

QY 1 GIGAVLXXXXVLTGTPALISWIKXXXXRK 31
||||| ||||||| :1
DB 1 GIGAVLK----VLTGTPALISWIKKKRQK 27

RESULT 3

US-08-449-741-12
Sequence 12, Application US/08449741
Patent No. 6022735

GENERAL INFORMATION:
APPLICANT: Curjel, David T.

APPLICANT: Birnstiel, Max L.
APPLICANT: Cotten, Matthew
APPLICANT: Wagner, Ernst
APPLICANT: Zatloukal, Kurt
APPLICANT: Plank, Christian
APPLICANT: Oberhauser, Berndt
APPLICANT: Schmidt, Walter G.M.
TITLE OF INVENTION: Composition for Introducing Nucleic Acid
TITLE OF INVENTION: Complexes Into Higher Eucaryotic Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Ave., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,741
FILING DATE: 25-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Plesner, Raz E.
REGISTRATION NUMBER: 34,331
REFERENCE/DOCKET NUMBER: 0652,0940007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: both
MOLECULE TYPE: peptide
US-08-449-741-12

Query Match 66.7%; Score 94; DB 3; Length 36;
Best Local Similarity 71.0%; Pred. No. 5.9e-10;
Matches 22; Conservative 1; Mismatches 4; Indels 4; Gaps 1;

QY 1 GIGAVLXXXXVLTGTPALISWIKXXXXRK 31
||||| ||||||| :1
DB 1 GIGAVLK----VLTGTPALISWIKKKRQK 27

RESULT 4

US-08-449-754-12
Sequence 12, Application US/08449754
Patent No. 6077663

GENERAL INFORMATION:
APPLICANT: Curjel, David T.
APPLICANT: Birnstiel, Max L.
APPLICANT: Cotten, Matthew
APPLICANT: Wagner, Ernst
APPLICANT: Zatloukal, Kurt
APPLICANT: Plank, Christian
APPLICANT: Oberhauser, Berndt
APPLICANT: Schmidt, Walter G.M.
TITLE OF INVENTION: Composition for Introducing Nucleic Acid
TITLE OF INVENTION: Complexes Into Higher Eucaryotic Cells
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.

ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,754
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/948,357
FILING DATE: 1992-09-23
ATTORNEY/AGENT INFORMATION:
NAME: Esmond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0652.0940004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: both
MOLECULE TYPE: peptide
US-08-449-754-12

Query Match 66.7%; Score 94; DB 3; Length 36;
Best Local Similarity 71.0%; Pred. No. 5.9e-10;
Matches 22: Conservative 1; Mismatches 4; Indels 4; Gaps 1;

OY 1 GIGAVLXXXXVLTGGLPALISMIRKXXXXRK 31
DB 1 GIGAVLK---VLTGGLPALISMIRKRRK 27

RESULT 5
US-09-465-646-12
Sequence 12, Application US/09465646
Patent No. 6274322
GENERAL INFORMATION:
APPLICANT: Curriel, David T.
APPLICANT: Birmstiel, Max L.
APPLICANT: Colten, Matthew
APPLICANT: Wagner, Ernst
APPLICANT: Zatloukal, Kurt
APPLICANT: Plank, Christian
APPLICANT: Oberhauser, Berndt
APPLICANT: Schmidt, Walter G.M.
TITLE OF INVENTION: COMPOSITION FOR INTRODUCING NUCLEIC ACID
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Ave., N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/465,646
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/449,754
FILING DATE: May 25, 1995

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/937,788
FILING DATE: September 2, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/864,759
FILING DATE: April 7, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/827,102
FILING DATE: January 30, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/767,788
FILING DATE: September 30, 1991
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/768,039
FILING DATE: September 30, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kamag, Andrea J.
REGISTRATION NUMBER: 43,703
REFERENCE/DOCKET NUMBER: 0652.0940009
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-6566
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: both
MOLECULE TYPE: peptide
US-09-465-646-12

Query Match 66.7%; Score 94; DB 4; Length 36;
Best Local Similarity 71.0%; Pred. No. 5.9e-10;
Matches 22: Conservative 1; Mismatches 4; Indels 4; Gaps 1;

OY 1 GIGAVLXXXXVLTGGLPALISMIRKXXXXRK 31
DB 1 GIGAVLK---VLTGGLPALISMIRKRRK 27

RESULT 6
US-07-643-343A-1
Sequence 1, Application US/07643343A
Patent No. 5235038
GENERAL INFORMATION:
APPLICANT: Blondelle, Sylvie E.
APPLICANT: Houghten, Richard A.
TITLE OF INVENTION: Deletion and Substitution
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cecchi & Stewart
STREET: 6 Becker Farm Road
CITY: Roseland
STATE: New Jersey
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: DM4.V2

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/643.343A
FILING DATE: 19910122
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Olstein, Elliot M.
REGISTRATION NUMBER: 24,025
REFERENCE/DOCKET NUMBER: 421250-139
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Apis mellifera
FEATURE:
NAME/KEY: mellitin peptide
PUBLICATION INFORMATION:
AUTHORS: Habermann, E.
TITLES: Jentsch, J.
TITLE: Sequenzanalyse des Mellitins aus
TITLES: den tryptischen und peptischen
TITLES: Spaltst oken.
JOURNAL: Hoppe-Seyler's Zeitschrift
VOLUME: 348
PAGES: 37-50
DATE: 1987
US-07-643-343A-1

Query Match 66.0%; Score 93; DB 1; Length 26;
Best Local Similarity 68.4%; Pred. No. 6.1e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

QY 1 GIGAVLKXXXVLTGTPALISWIKXXXXRXXRQ 38
DB 1 GIGAVLK---VLTGTPALISWIK---RK---RQ 26

RESULT 7
US-07-626-589-3
Sequence 3, Application US/07626589
Patent No. 5294698
GENERAL INFORMATION:
APPLICANT: Bomalski, John S.
APPLICANT: Clark, Robert
APPLICANT: Short, Robert
TITLE OF INVENTION: HUMAN PHOSPHOLIPASE ACTIVATING PROTEIN
TITLE OF INVENTION: AND METHODS FOR DIAGNOSIS OF RHEUMATOID ARTHRITIS
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESS: One Liberty Place - 46th floor
STREET: 1650 Market Street
CITY: Philadelphia
STATE: PA
COUNTRY: US
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/626.589
FILING DATE: 19901206

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hohenschutz, Liza D.
REGISTRATION NUMBER: 33,712
REFERENCE/DOCKET NUMBER: KOP-0001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 3
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-626-589-3

Query Match 66.0%; Score 93; DB 1; Length 26;
Best Local Similarity 68.4%; Pred. No. 6.1e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

QY 1 GIGAVLKXXXVLTGTPALISWIKXXXXRXXRQ 38
DB 1 GIGAVLK---VLTGTPALISWIK---RK---RQ 26

RESULT 8
US-08-079-445-7
Sequence 7, Application US/08079445
Patent No. 5440016
GENERAL INFORMATION:
APPLICANT: Blondelle, Sylvie E.
APPLICANT: Pinilla, Clemencia
APPLICANT: Eichler, Julia
APPLICANT: Houghten, Richard A.
TITLE OF INVENTION: PEPTIDES HAVING ANTI-MICROBIAL,
TITLE OF INVENTION: HEMOLYTIC, ANTI-MELLITIN OR ANTI-ENZYMATIC ACTIVITY
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/079.445
FILING DATE: 18-JUN-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-HP 9648
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 7
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 26
OTHER INFORMATION: /note- "C-terminal amino acid is
OTHER INFORMATION: amidated"
US-08-079-445-7

Query Match 66.0%; Score 93; DB 1; Length 26;
Best Local Similarity 68.4%; Pred. No. 6.1e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

QY 1 GIGAVLKXXXVLTGTPALISWIKXXXXXXRQO 38
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 GIGAVLK---VLTGTPALISWIK---RK---RQO 26

RESULT 9
US-07-943-709-58
; Sequence 58, Application US/07943709
; Patent No. 5556762

; GENERAL INFORMATION:
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Appel Jr., Jon R.
; APPLICANT: Blondelle, Silvie
; APPLICANT: Dooley, Colette T.
; APPLICANT: Elchler, Julia
; APPLICANT: Houghten, Richard A.
; TITLE OF INVENTION: SCANNING SYNTHETIC PEPTIDE COMBINATORIAL
; TITLE OF INVENTION: LIBRARIES: OLIGOPEPTIDE MIXTURE SETS HAVING ONE
; TITLE OF INVENTION: PREDETERMINED RESIDUE AT A SINGLE, PREDETERMINED POSITION,
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Dressler, Goldsmith, Shore, Suter &
; ADDRESSEE: Milamov, Ltd.
; STREET: 180 No. 5556762th Stetson, Suite 4700
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60601

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/943,709
; FILING DATE: 19920911
; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/797,551
; FILING DATE: 19-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Gamsen, Edward P.
; REGISTRATION NUMBER: 29,381
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5400
; TELEFAX: (312) 616-5460
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-943-709-58

Query Match 66.0%; Score 93; DB 1; Length 26;
Best Local Similarity 68.4%; Pred. No. 6.1e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

QY 1 GIGAVLKXXXVLTGTPALISWIKXXXXXXRQO 38
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 GIGAVLK---VLTGTPALISWIK---RK---RQO 26

RESULT 10
US-08-236-410-3
; Sequence 3, Application US/08236410
; Patent No. 5580968
; GENERAL INFORMATION:

; APPLICANT: Bomalaski, John S.
; APPLICANT: Clark, Mike A.
; APPLICANT: Short, Robert
; TITLE OF INVENTION: HUMAN PHOSPHOLIPASE ACTIVATING PROTEIN AND
; TITLE OF INVENTION: FOR DIAGNOSIS OF RHEUMATOID ARTHRITIS
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5580968-15
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103

; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk, 720 Kb
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/236,410
; FILING DATE: 02-MAY-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/147,925
; FILING DATE: 04-NOV-1993

; ATTORNEY/AGENT INFORMATION:
; NAME: Dianne B. Elderkin
; REGISTRATION NUMBER: 28,598
; REFERENCE/DOCKET NUMBER: MOPR-0028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-236-410-3

Query Match 66.0%; Score 93; DB 1; Length 26;
Best Local Similarity 68.4%; Pred. No. 6.1e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

QY 1 GIGAVLKXXXVLTGTPALISWIKXXXXXXRQO 38
| | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1 GIGAVLK---VLTGTPALISWIK---RK---RQO 26

RESULT 11
US-08-218-608-8
; Sequence 8, Application US/08218608
; Patent No. 5607859

; GENERAL INFORMATION:
; APPLICANT: BIEHMANN, KLAUS
; APPLICANT: JUHASZ, PETER
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR MASS
; TITLE OF INVENTION: SPECTROPHETIC MOLECULAR WEIGHT DETERMINATION OF POLYIONIC
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02210

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25

```

1 CURRENT APPLICATION DATA:
2 APPLICATION NUMBER: US/08/218,608
3 FILING DATE: 28-MAR-1994
4 CLASSIFICATION: 436
5 ATTORNEY/AGENT INFORMATION:
6 NAME: GATES, EDWARD R.
7 REGISTRATION NUMBER: 31,616
8 REFERENCE/DOCKET NUMBER: M0656/7013
9 TELECOMMUNICATION INFORMATION:
10 TELEPHONE: 617-720-3500
11 TELEFAX: 617-720-2441
12 INFORMATION FOR SEQ ID NO: 8:
13 SEQUENCE CHARACTERISTICS:
14 LENGTH: 26 amino acids
15 TYPE: amino acid
16 STRANDEDNESS: single
17 TOPOLOGY: linear
18 MOLECULE TYPE: peptide
19 HYPOTHETICAL: NO
20 ANTI-SENSE: NO
21 ORIGINAL SOURCE:
22 ORGANISM: Apis mellifera
23
24 US-08-218-608-8

```

Query Match	66.08;	Score 93;	DB 1;	length 26;
Best Local Similarity	68.48;	Pred. No.	6.1e-10;	
Matches	26;	Conservative	0;	Indels 12;
		Mismatches	0;	Gaps 3

```

Qy      1 GIGANLKKXXXXVLTTGTPALISWIKXXXXRKRXXXXRQD 38
        |||||  |||||  |||
Db      1 GIGANLK----VLTGTGTPALISWIK----RK----RQD 26

```

RESULT 12
 US-08-434-1120-114
 : Sequence 114, Application US/08434120
 : Patent No. 563479
 :
 : GENERAL INFORMATION:
 : APPLICANT: Baker, Margaret A.
 : APPLICANT: Jacob, Leonard S.
 : APPLICANT: Malloy, W. Lee
 :
 : TITLE OF INVENTION: Treatment of Gynecological
 : TITLE OF INVENTION: Malignancies with
 : TITLE OF INVENTION: Biologically Active Peptides
 : NUMBER OF SEQUENCES: 117
 :
 : CORRESPONDENCE ADDRESSES:
 : ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
 : ADDRESSEE: Cecchi & Stewart
 : STREET: 6 Becker Farm Road
 : CITY: Roseland
 : STATE: New Jersey
 : COUNTRY: USA
 : ZIP: 07068
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5 Inch diskette
 : COMPUTER: IBM PS/2
 : OPERATING SYSTEM: PC-DOS
 : SOFTWARE: DW4.V2
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/434,120
 : FILING DATE:
 :
 : CLASSIFICATION: 514
 :
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/08/297,950
 : FILING DATE:
 :
 : APPLICATION NUMBER: US/08/226,108
 : FILING DATE:
 :
 : APPLICATION NUMBER: US/07/937,462
 : FILING DATE:
 :
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Olstein, Elliot M.
 : REGISTRATION NUMBER: 24,025
 : REFERENCE/DOCKET NUMBER: 421250-194

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 114:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Apis mellifera
FEATURE:
NAME/KEY: mellitin peptide
PUBLICATION INFORMATION:
AUTHORS: Habermann, E.
AUTHORS: Jentsch, J.
TITLE: Sequenzanalyse des Mellitins aus
        Spaltstücken
JOURNAL: Hoppe-Seyler's Zeitschrift
        für
        Journal: Physiol. Chem.
VOLUME: 348
PAGES: 1987
DATE: 1987
US-08-434-120-114

```

Query Match	66.0%	Score	93	DB	1	Length	26
Best Local Similarity	68.4%	Pred. NO.	6.1e-10				
Matches	26	Conservative	0	Mismatches	0	Indels	12
						Gaps	3

```
QY      I GIGAVLXXXXVLTGTGLPALISWIKXXXXRKRXXXXRQQ 38
        ||||| . ||||||||| | ||
Db       1 GIGAVLK----VLTTGTPALISWIK----RK-----RQQ 26
```

RESULT 13
US-08-465-325-93
Sequence 93, Application US/08465325
Patent No. 5686563
GENERAL INFORMATION:
APPLICANT: Magalain Pharmaceuticals Inc.
APPLICANT: 5110 Campus Drive
APPLICANT: Plymouth Meeting, PA 19462
TITLE OF INVENTION: Biologically Active Peptides Having
TITLE OF INVENTION: N-Terminal Substitutions
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Flinegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I. Street, N.W. Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,325
FILING DATE: 05-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/184,462
FILING DATE: 18-JAN-94
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/891,201
FILING DATE: 01-JUN-92
ATTORNEY/AGENT INFORMATION:
NAME: Fordis, Jean B
REGISTRATION NUMBER: 32,984
REFERENCE/DOCKET NUMBER: 05387.0021-03000

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 93:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-465-325-93

Query Match 66.0%; Score 93; DB 1; Length 26;
Best Local Similarity 68.4%; Pred. No. 6.1e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

QY 1 GIGAVLXXXXVLTGTPALISWIKXXXXRXXXXRQ 38
DB 1 GIGAVLK-----VLTGTPALISWIK-----RK-----RQ 26

RESULT 14

US-08-434-761-7
Sequence 7, Application US/08434761
Patent No. 5698673
GENERAL INFORMATION:
APPLICANT: Biondelle, Sylvie E.
APPLICANT: Pinilla, Clemencia
APPLICANT: Eichler, Jutta
APPLICANT: Houghten, Richard A.
TITLE OF INVENTION: PEPTIDES HAVING ANTI-MICROBIAL,
TITLE OF INVENTION: HEMOLYTIC, ANTI-MELITTIN OR ANTI-ENZYMATIC ACTIVITY
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: CAMPBELL AND FLORES
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States of America
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,761
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/079,445
FILING DATE: 18-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-HP 9648
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 26
OTHER INFORMATION: /note- "C-terminal amino acid is
OTHER INFORMATION: amended"
US-08-434-761-7

Query Match 66.0%; Score 93; DB 1; Length 26;
Best Local Similarity 68.4%; Pred. No. 6.1e-10;

Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;
QY 1 GIGAVLXXXXVLTGTPALISWIKXXXXRXXXXRQ 38
DB 1 GIGAVLK-----VLTGTPALISWIK-----RK-----RQ 26

RESULT 15

US-08-039-557-5
Sequence 5, Application US/08039557
Patent No. 5714467
GENERAL INFORMATION:
APPLICANT: Boman, Hans G.
APPLICANT: Merrifield, Robert B.
APPLICANT: Andrew, David
TITLE OF INVENTION: ANTIBACTERIAL AND ANTIMALARIAL HYBRID
TITLE OF INVENTION: PEPTIDES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/039,557
FILING DATE: 26-MAR-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,462
FILING DATE: 05-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/449,593
FILING DATE: 12-DEC-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/336,777
FILING DATE: 12-APR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 600-1-110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 26 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Peptide
HYPOTHETICAL: YES
FRAGMENT TYPE: N-terminal
US-08-039-557-5

Query Match 66.0%; Score 93; DB 1; Length 26;
Best Local Similarity 68.4%; Pred. No. 6.1e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

QY 1 GIGAVLXXXXVLTGTPALISWIKXXXXRXXXXRQ 38
DB 1 GIGAVLK-----VLTGTPALISWIK-----RK-----RQ 26

Search completed: May 9, 2003, 16:19:27
Job time: 14.7843 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:15:46 ; Search time 16.0196 Seconds
(Without alignments)
218.294 Million cell updates/sec

Title: US-09-851-422a-2
Perfect score: 141
Sequence: 1 GIGAVLKXXXXVLTGHPALISWIKXXXXRRXXXXRQ 38

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB pep:.*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB pep:.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB pep:.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB pep:.*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:.*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB pep:.*
7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB pep:.*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB pep:.*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB pep:.*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB pep:.*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB pep:.*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB pep:.*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB pep:.*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB pep:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	66.0	26	9 US-09-999-745-50	Sequence 50, Appl
2	93	66.0	26	9 US-09-554-000-34	Sequence 34, Appl
3	93	66.0	26	9 US-09-820-053A-72	Sequence 72, Appl
4	93	66.0	26	10 US-09-030-619-126	Sequence 26, Appl
5	93	66.0	70	9 US-08-847-208-26	Sequence 26, Appl
6	92	65.2	26	10 US-09-752-533-1	Sequence 1, Appl
7	87.5	62.1	27	10 US-09-917-940-25	Sequence 25, Appl
8	72	51.1	26	10 US-09-030-619-120	Sequence 120, Appl
9	64	45.4	28	10 US-09-909-652-7	Sequence 7, Appl
10	64	45.4	28	10 US-09-030-619-121	Sequence 121, Appl
11	63	44.7	15	10 US-09-756-250B-3	Sequence 3, Appl
12	51	36.2	26	10 US-09-030-619-122	Sequence 122, Appl
13	51	36.2	26	10 US-09-030-619-124	Sequence 124, Appl
14	51	36.2	26	10 US-09-030-619-128	Sequence 128, Appl
15	50	35.5	26	9 US-09-030-619-130	Sequence 130, Appl
16	49	34.8	23	9 US-08-820-053A-69	Sequence 69, Appl
17	48	34.0	13	10 US-09-756-250B-2	Sequence 2, Appl
18	47	33.3	1603	9 US-10-208-948-2	Sequence 24, Appl
19	47	33.3	2159	9 US-10-208-948-24	Sequence 24, Appl

20	45	31.9	26	10 US-09-030-619-125	Sequence 125, Appl
21	43	30.5	26	10 US-09-030-619-123	Sequence 123, Appl
22	43	30.5	425	10 US-09-798-029-2	Sequence 2, Appl
23	42	29.8	110	10 US-09-798-029-18	Sequence 18, Appl
24	42	29.8	532	9 US-09-941-947A-38	Sequence 38, Appl
25	41	29.1	279	12 US-10-014-927-19	Sequence 19, Appl
26	41	29.1	529	9 US-10-051-902-28	Sequence 28, Appl
27	41	29.1	529	9 US-10-051-909-28	Sequence 28, Appl
28	40.5	28.7	761	9 US-09-147-947-4	Sequence 4, Appl
29	40	28.4	345	10 US-09-943-671-3	Sequence 3, Appl
30	40	28.4	4150	9 US-09-808-880-2	Sequence 2, Appl
31	39.5	28.0	416	10 US-09-927-616A-4	Sequence 4, Appl
32	39.5	28.0	475	9 US-09-738-626-5149	Sequence 5149, Appl
33	39	27.7	197	10 US-09-815-242-13736	Sequence 13736, A
34	39	27.7	316	10 US-09-115-150-5	Sequence 5, Appl
35	39	27.7	389	9 US-10-081-816-21	Sequence 21, Appl
36	39	27.7	539	9 US-10-051-902-26	Sequence 26, Appl
37	39	27.7	539	9 US-10-051-909-26	Sequence 26, Appl
38	38.5	27.3	663	10 US-09-991-630-25	Sequence 25, Appl
39	38.5	27.3	2597	10 US-09-905-129-2	Sequence 10, Appl
40	38.5	27.3	2597	10 US-09-905-129-10	Sequence 10, Appl
41	38.5	27.3	2597	10 US-09-905-129-13	Sequence 13, Appl
42	38.5	27.3	2597	10 US-09-991-630-12	Sequence 2, Appl
43	38.5	27.3	2597	10 US-09-991-630-10	Sequence 10, Appl
44	38.5	27.3	2597	10 US-09-991-630-13	Sequence 13, Appl
45	38	27.0	17	10 US-09-859-377-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1

US-09-999-745-50
Sequence 50, Application US/09999745
Patent No. US20020157120A1
GENERAL INFORMATION:
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
APPLICANT: Baird, Geoffrey
TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
FILE REFERENCE: RESENT470-1
CURRENT APPLICATION NUMBER: US/09/999,745
CURRENT FILING DATE: 2001-10-23
PRIOR APPLICATION NUMBER: 09/316,920
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.0
SEQ ID NO 50
LENGTH: 26
TYPE: PRT
ORGANISM: Apis mellifera
US-09-999-745-50

Query Match

66.0%; Score 93; DB 9; Length 26;

Best Local Similarity 68.4%; Pred. No. 4.9e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

Query 1 GIGAVLKXXXXVLTGHPALISWIKXXXXRRXXXXRQ 38
Db 1 GIGAVLK----VLTGHPALISWIK----RK----RQ 26

RESULT 2

US-09-554-000-34
Sequence 34, Application US/09554000
Patent No. US20020165364A1
GENERAL INFORMATION:
APPLICANT: Tsien, Roger Y.
APPLICANT: Miyawaki, Atsushi
TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
TITLE OF INVENTION: DETECTION OF ANALYTES
FILE REFERENCE: 07257/042001
CURRENT APPLICATION NUMBER: US/09/554,000

;; CURRENT FILING DATE: 2000-04-20
;; PRIOR APPLICATION NUMBER: 08/818,252
;; PRIOR FILING DATE: 1997-03-14
;; NUMBER OF SEQ ID NOS: 56
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 34
;; LENGTH: 26
;; TYPE: PRT
;; ORGANISM: Apis mellifera
US-09-554-000-34

Query Match 66.0%; Score 93; DB 9; Length 26;
Best Local Similarity 68.4%; Pred. No. 4.9e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

Oy 1 GIGAVLKKXXXVLTGTPALISWIKXXXXRRXXXXXROQ 38
Db 1 GIGAVLK----VLTGTPALISWIK----RK----ROQ 26

RESULT 3

US-09-820-053A-72
;; Sequence 72, Application US/09820053A
;; Publication No. US2003008243A1
;; GENERAL INFORMATION:
;; APPLICANT: Owen, Donald R.
;; TITLE OF INVENTION: SHORT BIOACTIVE PEPTIDES
;; FILE REFERENCE: HELX027
;; CURRENT APPLICATION NUMBER: US/09/820,053A
;; CURRENT FILING DATE: 2001-03-28
;; NUMBER OF SEQ ID NOS: 165
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 72
;; LENGTH: 26
;; TYPE: PRT
;; ORGANISM: ARTIFICIAL SEQUENCE
;; FEATURE:
;; OTHER INFORMATION: SYNTHETIC SEQUENCE
;; NAME/KEY: MOD-RES
;; LOCATION: (26)
;; OTHER INFORMATION: AMIDATION
US-09-820-053A-72

Query Match 66.0%; Score 93; DB 9; Length 26;
Best Local Similarity 68.4%; Pred. No. 4.9e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

Oy 1 GIGAVLKKXXXVLTGTPALISWIKXXXXRRXXXXXROQ 38
Db 1 GIGAVLK----VLTGTPALISWIK----RK----ROQ 26

RESULT 4

US-09-030-619-216
;; Sequence 216, Application US/09030619B
;; Patent No. US20020035061A1
;; GENERAL INFORMATION:
;; APPLICANT: Krieger, Timothy J.
;; APPLICANT: Taylor, Robert
;; APPLICANT: Erfile, Douglas
;; APPLICANT: Fraser, Janet R.
;; APPLICANT: West, Michael H.P.
;; APPLICANT: McNicol, Patricia J.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
;; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
;; TITLE OF INVENTION: WITH ANTIBIOTICS
;; FILE REFERENCE: 660081.406
;; CURRENT APPLICATION NUMBER: US/09/030,619B
;; CURRENT FILING DATE: 1998-02-25
;; NUMBER OF SEQ ID NOS: 232
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 216
;; LENGTH: 26

;; TYPE: PRT
;; ORGANISM: Apis mellifera
US-09-030-619-216

Query Match 66.0%; Score 93; DB 10; Length 26;
Best Local Similarity 68.4%; Pred. No. 4.9e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

Oy 1 GIGAVLKKXXXVLTGTPALISWIKXXXXRRXXXXXROQ 38
Db 1 GIGAVLK----VLTGTPALISWIK----RK----ROQ 26

RESULT 5

US-09-847-208-26
;; Sequence 26, Application US/09847208
;; Publication No. US20030082190A1
;; GENERAL INFORMATION:
;; APPLICANT: Saxon, Andrew
;; APPLICANT: Zhang, Ke
;; APPLICANT: Zhu, Daocheng
;; TITLE OF INVENTION: FUSION MOLECULES AND TREATMENT OF
;; TITLE OF INVENTION: IGE-MEDIATED ALLERGIC DISEASES
;; FILE REFERENCE: UC67,002A
;; CURRENT APPLICATION NUMBER: US/09/847,208
;; CURRENT FILING DATE: 2001-05-01
;; NUMBER OF SEQ ID NOS: 177
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 26
;; LENGTH: 70
;; TYPE: PRT
;; ORGANISM: Apis mellifera(Honeybee)Apis cerana(Ind. honeybee)
US-09-847-208-26

Query Match 66.0%; Score 93; DB 9; Length 70;
Best Local Similarity 68.4%; Pred. No. 1.4e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

Oy 1 GIGAVLKKXXXVLTGTPALISWIKXXXXRRXXXXXROQ 38
Db 44 GIGAVLK----VLTGTPALISWIK----RK----ROQ 69

RESULT 6

US-09-752-533-1
;; Sequence 1, Application US/09752533
;; Patent No. US20020107389A1
;; GENERAL INFORMATION:
;; APPLICANT: Coutts, Stephen M.
;; APPLICANT: Jones, David S.
;; APPLICANT: Livingston, Douglas A.
;; APPLICANT: Yu, Lin
;; TITLE OF INVENTION: CONJUGATES OF CHEMICALLY DEFINED
;; TITLE OF INVENTION: NON-POLYMERIC VALENCY PLATFORM MOLECULES AND BIOLOGICALLY
;; TITLE OF INVENTION: ACTIVE MOLECULES
;; FILE REFERENCE: 252312005704
;; CURRENT APPLICATION NUMBER: US/09/752,533
;; CURRENT FILING DATE: 2000-12-29
;; PRIOR APPLICATION NUMBER: 08/769,041
;; PRIOR FILING DATE: 1996-12-18
;; PRIOR APPLICATION NUMBER: 08/453,254
;; PRIOR FILING DATE: 1995-05-30
;; PRIOR APPLICATION NUMBER: 08/152,506
;; PRIOR FILING DATE: 1993-11-15
;; PRIOR APPLICATION NUMBER: 07/914,869
;; PRIOR FILING DATE: 1992-07-15
;; PRIOR APPLICATION NUMBER: 08/118,055
;; PRIOR FILING DATE: 1993-09-08
;; NUMBER OF SEQ ID NOS: 9
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 1
;; LENGTH: 26
;; TYPE: PRT

ORGANISM: Apis cerana
FEATURE:
NAME/KEY: AMIDATION
LOCATION: 26
OTHER INFORMATION: Gln attached to an Amide
US-09-752-533-1

Query Match 65.2%; Score 92; DB 10; Length 26;
Best Local Similarity 84.0%; Pred. No. 7.2e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 GIGAVLKXXXVLTGTPALISMIK 25
DB 1 GIGAVLK---VLTGTPALISMIK 21

RESULT 7
US-09-917-340-25
Sequence 25, Application US/09917340
Patent No. US20020090369A1
GENERAL INFORMATION:
APPLICANT: Murphy, Christopher J.
APPLICANT: McAnulty, Jonathan F.
APPLICANT: Reid, Ted W.
TITLE OF INVENTION: Transplant Media
FILE REFERENCE: TPLANT-06468
CURRENT APPLICATION NUMBER: US/09/917,340
CURRENT FILING DATE: 2001-07-29
PRIOR APPLICATION NUMBER: 60/221,632
PRIOR FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: 60/249,602
PRIOR FILING DATE: 2000-11-17
PRIOR APPLICATION NUMBER: 60/290,932
PRIOR FILING DATE: 2001-05-15
NUMBER OF SEQ ID NOS: 96
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 25
LENGTH: 27
TYPE: PRT
ORGANISM: Apis mellifera
US-09-917-340-25

Query Match 62.1%; Score 87.5; DB 10; Length 27;
Best Local Similarity 65.8%; Pred. No. 4.3e-08;
Matches 25; Conservative 0; Mismatches 2; Indels 11; Gaps 3;

QY 1 GIGAVLKXXXVLTGTPALISMIKXXXKXXNRQ 38
DB 1 GIGAVLK---VLTGTPALISMI---SRK---KROQ 27

RESULT 8
US-09-030-619-120
Sequence 120, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 120
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-120

Query Match 51.1%; Score 72; DB 10; Length 26;
Best Local Similarity 81.8%; Pred. No. 1.7e-05;
Matches 18; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 GIGAVLKXXXVLTGTPALIS 22
DB 9 GIGAVLK---VLTGTPALIS 26

RESULT 9
US-09-909-652-7
Sequence 7, Application US/09909652
Patent No. US20020025537A1
GENERAL INFORMATION:
APPLICANT: Kairos Scientific, Inc.
APPLICANT: Bylina, Edward J.
APPLICANT: Coleman, William J.
APPLICANT: Youvan, Douglas C.
TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR GENERATING
TITLE OF INVENTION: AND SCREENING COMPOUNDS THAT AFFECT CELL VIABILITY
FILE REFERENCE: 22346-7001
CURRENT APPLICATION NUMBER: US/09/909,652
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: US 60/219,179
PRIOR FILING DATE: 2000-07-19
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: CEMA analogue of cecropin-melittin hybrid
US-09-909-652-7

Query Match 45.4%; Score 64; DB 10; Length 28;
Best Local Similarity 80.0%; Pred. No. 0.00042;
Matches 16; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 GIGAVLKXXXVLTGTPAL 20
DB 9 GIGAVLK---VLTGTPAL 24

RESULT 10
US-09-030-619-121
Sequence 121, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 121
LENGTH: 28
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-121

Query Match 45.4%; Score 64; DB 10; Length 28;
Best Local Similarity 80.0%; Pred. No. 0.00042;
Matches 16; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 1 GIGAVLXXXXVLTGGLPAL 20
DB 9 GIGAVLXXXXVLTGGLPAL 24

RESULT 11
US-09-756-250B-3
Sequence 3, Application US/09756250B
Patent No. US20020058020A1
GENERAL INFORMATION:
APPLICANT: No. US20020058020A1artis AG
TITLE OF INVENTION: Isolated amphiphilic peptides derived from the cytoplasmic tail of
FILE OF INVENTION: envelope proteins
FILE REFERENCE: 4-30663a/C1/CP1 2042
CURRENT APPLICATION NUMBER: US/09/756,250B
CURRENT FILING DATE: 2001-12-13
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patent version 3.0
SEQ ID NO 3
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Shortened analogue of melittin peptide
NAME/KEY: MUTAGEN
LOCATION: (1)..(15)
US-09-756-250B-3

Query Match 44.7%; Score 63; DB 10; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.00031;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 VLTGGLPALISWT 24
DB 3 VLTGGLPALISWT 15

RESULT 12
US-09-030-619-122
Sequence 122, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 122
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-122

Query Match 36.2%; Score 51; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.06;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 VLTGGLPALIS 22
DB 16 VLTGGLPALIS 26

RESULT 13
US-09-030-619-124
Sequence 124, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE OF INVENTION: WITH ANTIBIOTICS
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 124
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Cationic Peptide Analogue
US-09-030-619-124

Query Match 36.2%; Score 51; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VLTGGLPALIS 22
DB 16 VLTGGLPALIS 26

RESULT 14
US-09-030-619-128
Sequence 128, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: McNicol, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 128
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fusion Peptides
NAME/KEY: VARIANT
LOCATION: (4)...(4)
OTHER INFORMATION: Xaa - Hydrophilic Residue
LOCATION: (5)...(6)

OTHER INFORMATION: Xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (7)...(8)
OTHER INFORMATION: Xaa - Hydrophilic Residue
NAME/KEY: VARIANT
LOCATION: (9)...(9)
OTHER INFORMATION: Xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (10)...(11)
OTHER INFORMATION: Xaa - Hydrophilic Residue
NAME/KEY: VARIANT
LOCATION: (12)...(13)
OTHER INFORMATION: Xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (14)...(15)
OTHER INFORMATION: Xaa - Hydrophilic Residue
US-09-030-619-128

Query Match 36.2%; Score 51; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 0.06;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 VLTGGLPALIS 22
DB 16 VLTGGLPALIS 26

RESULT 15

US-09-030-619-130
Sequence 130, Application US/09030619B
Patent No. US20020035061A1
GENERAL INFORMATION:
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Monicoll, Patricia J.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
FILE REFERENCE: 660081.406
CURRENT APPLICATION NUMBER: US/09/030,619B
CURRENT FILING DATE: 1998-02-25
NUMBER OF SEQ ID NOS: 232
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 130
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Fusion Peptides
NAME/KEY: VARIANT
LOCATION: (4)...(4)
OTHER INFORMATION: Xaa - Hydrophilic Residue
NAME/KEY: VARIANT
LOCATION: (5)...(6)
OTHER INFORMATION: Xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (7)...(8)
OTHER INFORMATION: Xaa - Hydrophilic Residue
NAME/KEY: VARIANT
LOCATION: (9)...(9)
OTHER INFORMATION: Xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (10)...(11)
OTHER INFORMATION: Xaa - Hydrophilic Residue
NAME/KEY: VARIANT
LOCATION: (12)...(13)
OTHER INFORMATION: Xaa - Hydrophobic Residue
NAME/KEY: VARIANT
LOCATION: (14)...(15)
OTHER INFORMATION: Xaa - Hydrophilic Residue

US-09-030-619-130

Query Match 35.5%; Score 50; DB 10; Length 26;
Best Local Similarity 90.9%; Pred. No. 0.089;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 12 VLTGGLPALIS 22
DB 16 ILTTGLPALIS 26

Search completed: May 9, 2003, 16:20:17
Job time: 17.0196 secs

THIS PAGE BLANK (U8FT0)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:12:15 ; Search time 23.8431 seconds
(without alignments)
153.214 Million cell updates/sec

Title: US-09-851-422a-2
Perfect score: 141
Sequence: 1 GIGAVLXXXXXVLTGTPALISWIKXXXXXRRQ 38

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	66.0	26	1 MEHBCI	melittin, major -
2	93	66.0	70	1 MEHBI	melittin, major, p
3	87.5	62.1	27	1 MEHB2	melittin, minor -
4	87	61.7	26	1 MEHBCD	melittin - giant h
5	82	58.2	26	1 MEHBCF	melittin - little
6	46	32.6	642	2 I46534	renal sodium-depen
7	45	31.9	197	2 A10115	conserved hypotet
8	45	31.9	357	2 S11137	class I histocompa
9	45	31.9	357	2 S11134	class I histocompa
10	45	31.9	357	2 S11135	class I histocompa
11	45	31.9	357	2 S11140	class I histocompa
12	45	31.9	357	2 S11139	class I histocompa
13	45	31.9	357	2 T72218	class I histocompa
14	44	31.2	357	2 S11133	class I histocompa
15	44	31.2	365	2 I56053	class I histocompa
16	44	31.2	714	2 T25056	hypothetical prote
17	44	31.2	749	2 A75506	conserved hypotet
18	44	31.2	762	2 H87466	beta-D-glucosidase
19	44	31.2	4344	1 A53489	dyein heavy chain
20	43.5	30.9	200	2 G75550	conserved hypotet
21	43	30.5	174	2 G95083	conserved domain p
22	43	30.5	174	2 C97951	hypothetical prote
23	43	30.5	283	2 T44504	hypothetical prote
24	43	30.5	303	2 F86158	alternative splici
25	43	30.5	462	2 E91156	probable amino aci
26	43	30.5	462	2 B86002	probable amino aci
27	43	30.5	462	2 E65131	hypothetical 47.5
28	43	30.5	744	2 AG1382	transport protein
29	42.5	30.1	330	2 C87036	probable conserved

30	42.5	30.1	1105	2 S76557	cardamyl-phosphat
31	42	29.8	239	2 AE1781	hypothetical prote
32	42	29.8	262	2 E90498	conserved hypotet
33	42	29.8	478	2 D83043	conserved hypotet
34	42	29.8	639	2 B48189	sodium/phosphate c
35	42	29.8	667	2 T23010	hypothetical prote
36	41	29.1	198	2 T47402	nodulin-like prote
37	41	29.1	237	2 E86223	hypothetical prote
38	41	29.1	242	2 T45421	probable 2-hydroxy
39	41	29.1	262	2 B96937	chemotaxis motilit
40	41	29.1	315	2 AB1577	cobalamine synthe
41	41	29.1	346	1 JH0387	alkanal monooxygen
42	41	29.1	354	2 JH0413	alkanal monooxygen
43	41	29.1	357	2 S11136	class I histocompa
44	41	29.1	359	1 H1H112	class I histocompa
45	41	29.1	364	2 I72217	class I histocompa

ALIGNMENTS

RESULT 1
MEHBCI
melittin, major - Indian honeybee
C:Species: Apis mellifera cerana (Indian honeybee)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 15-Oct-1996
C:Accession: A01762

R:Krell, G., 33, 241-244, 1973
FEBS Lett. 33, 241-244, 1973
A:Title: Structure of melittin isolated from two species of honey bees.

A:Reference number: A91392

A:Accession: A01762

A:Molecule type: protein

A:Residues: 1-26 <KRD>

A>Note: the Indian bee sequence is identical with that of honeybee (Apis mellifera) a

C:Superfamily: melittin major

C:Keywords: amidated carboxyl end; hemolysis; homotetramer; venom

F:26/Modified site: amidated carboxyl end (Gln) #status experimental

Query Match 66.0%; Score 93; DB 1; Length 26;

Best Local Similarity 68.4%; Pred. No. 7.1e-09;

Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

OY 1 GIGAVLXXXXXVLTGTPALISWIKXXXXXRRQ 38
DB 1 GIGAVLK-----VLTGTPALISWIK---RK---RQ 26

RESULT 2
MPHBI
melittin, major, precursor [validated] - honeybee

N:Contains: melittin F
C:Species: Apis mellifera (honeybee)
C:Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 15-Sep-2000
C:Accession: A91133; A91640; B61285; S23131; A01761; A18880

R:Vlasak, R.; Unger-Ullmann, C.; Krell, G.; Fritschauf, A.M.
Eur. J. Biochem. 135, 123-126, 1983

A:Title: Nucleotide sequence of cloned cDNA coding for honeybee prepro-melittin.
A:Reference number: A91133; MVID:83287387; PMID:6309516

A:Accession: A91133

A:Molecule type: mRNA

A:Residues: 1-70 <VLAV>

A:Cross-references: GB:X02007; NID:95621; PIDD:CA26038.1; PIDD:95622

R:Hebermann, E.; Jentsch, J.
Hoppe-Seyler's Z. Physiol. Chem. 348, 37-50, 1967

A:Title: Sequenzanalyse des Melittins aus den tryptischen und peptischen Spaltstuecke

A:Reference number: A91640; MVID:68337913; PMID:5592400

A:Accession: A91640

A:Molecule type: protein

A:Residues: 44-69 <HAB>

R:Gauldie, J.; Hanson, J.M.; Shipollini, R.A.; Vernon, C.A.
Eur. J. Biochem. 83, 405-410, 1978

A:Title: The structures of some peptides from bee venom.

```

A:Reference number: A91253; MUID:7812668; PMID:631126
A:Accession: B61285
A:Molecule type: protein
A:Residues: 51-69 <GAU>
R:Ramalingam, K.; Bello, J.
Biochem. J. 284, 663-665, 1992
A:Title: Effect of permethylation on the haemolytic activity of mellitin.
A:Reference number: 523131; MUID:92321983; PMID:1622387
A:Accession: S23131
A:Status: preliminary
A:Molecule type: protein
A:Residues: 44-69 <RAM>
R:Schröder, E.; Luebke, K.; Lehmann, M.; Beetz, I.
Experientia 27, 764-765, 1971
A:Title: Haemolytic activity and action on the surface tension of aqueous solutions of mellitin.
A:Reference number: A91266; MUID:72096668; PMID:5133482
A:Contents: annotation; synthesis
A>Note: The structure of mellitin was confirmed by synthesis of a peptide with full hemolytic activity.
Biochem. Biophys. Res. Commun. 27, 275-280, 1967
A:Title: The isolation of N-formylglycine from a polypeptide present in bee venom.
A:Reference number: A90165; MUID:67248282; PMID:6040373
A:Contents: annotation
A>Note: About 10% of mellitin is formylated at the amino end
R:Luebke, K.; Matthes, S.; Kloss, G.
Experientia 27, 765-767, 1971
A:Title: Isolation and structure of N(alpha)-formyl mellitin.
A:Reference number: A91267; MUID:72096669; PMID:5139483
A:Contents: annotation; synthesis
A>Note: N-formyl-mellitin major was isolated from the venom and its structure was confirmed by synthesis.
R:Mollay, C.; Vilas, U.; Krell, G.
Proc. Natl. Acad. Sci. U.S.A. 79, 2260-2263, 1982
A:Title: Cleavage of honeybee prepomellitin by an endoprotease from rat liver microsomes.
A:Reference number: A18860; MUID:82247982; PMID:7048315
A:Contents: annotation; intact signal sequence after cleavage by partially purified signal peptidase.
R:Eisenberg, D.; Grisham, M.; Terwilliger, T.C.
submitted to the Brookhaven Protein Data Bank, October 1990
A:Reference number: A50496; PDB:2MT
A:Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 44-69
R:Terwilliger, T.C.; Eisenberg, D.
J. Biol. Chem. 257, 6010-6015, 1982
A:Title: The structure of mellitin. I. Structure determination and partial refinement.
A:Reference number: A30639; MUID:82189958; PMID:7076661
A:Contents: annotation; X-ray crystallography, 2.0 angstroms
R:Terwilliger, T.C.; Eisenberg, D.
J. Biol. Chem. 257, 6016-6022, 1982
A:Title: The structure of mellitin. II. Interpretation of the structure.
A:Reference number: A30640; MUID:82189959; PMID:7076662
A:Contents: annotation; X-ray crystallography, 2.0 angstroms
C:Keywords: amidated carboxyl end; blocked amino end; hemolysis; homotetramer; venom
C:Keywords: amidated carboxyl end; blocked amino end; hemolysis; homotetramer; venom
P:1-21/Domains: signal sequence #status experimental <SIG>
P:22-43/Domains: propeptide #status experimental <PRO>
P:44-69/Product: mellitin major #status experimental <MAJ>
P:51-69/Product: mellitin F #status experimental <MLF>
P:44/Modified site: formylated amino end (Gly) (in mature form) (partial) #status experimental
P:69/Modified site: amidated carboxyl end (Gln) (amide in mature form from following gly)

Query Match 66.0%; Score 93; DB 1; Length 70;
Best Local Similarity 68.4%; Pred. NO. 2e-08;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;

Oy 1 GIGAVLKKXXVLTGTALPALSIMIKXXXXRXXRQ 38
    ||||| ||||||||| |||
Db 44 GIGAVLK-----VLTGTALPALSIMIK-----RK-----RQ 69

RESULT 3
MEH32
mellitin, minor - honeybee
C:Species: Apis mellifera (honeybee)
C:Date: 13-Jul-1981 #sequence_revision 13-Jul-1981 #text_change 15-Oct-1996
C:Accession: A01764

```

```

R: Habermann, I.; Jentsch, J.
Hoppe-Seyler's Z. Physiol. Chem. 348, 37-50, 1967
A: Title: Sequenzanalyse des Mellitins aus den tryptischen und peptischen Spaltstuecke
A: Reference number: A91640; MUID: 68327913; PMID: 5592400
A: Accession: A01764
A: Molecule type: protein
A: Residues: 1-27 <HAB>
R: Schroeder, E.; Luebke, K.; Lehmann, M.; Beetz, I.
Experientia 27, 764-765, 1971
A: Title: Hemolytic activity and action on the surface tension of aqueous solutions of
A: Reference number: A91266; MUID: 72098668; PMID: 5139482
A: Contents: annotation; synthesis
A: Note: the structure was confirmed by synthesis of a peptide with full hemolytic act
C: Superfamily: mellitin major
C: Keywords: amidated carboxyl end; hemolysis; homotetramer; venom
F: 27/Modified site: amidated carboxyl end (Gln) #status experimental

Query Match 62.1%; Score 87.5; DB 1; Length 27;
Best Local Similarity 65.8%; Pred. No. 6.1e-08;
Matches 25; Conservative 0; Mismatches 2; Indels 11; Gaps 3;

Oy 1 GIGAVLXXXXXVLTGTGALPALSMTKXXXXRRXXXXRQ 38
Db 1 GIGAVLK----VLTGTGALPALSMT---SRK---KRQ 27
||||| ||||||||| || |||

RESULT 4
MEHBCD
mellitin - giant honeybee (tentative sequence)
C: Species: Apis mellifera dorsata (giant honeybee)
C: Date: 22-Jun-1981 #sequence_revision 22-Jun-1981 #text_change 31-Mar-2000
C: Accession: A01763
R: Kreil, G.
FEBS Lett. 54, 100-102, 1975
A: Title: The structure of Apis dorsata mellitin: phylogenetic relationships between h
A: Reference number: A01763; MUID: 75168194; PMID: 1093875
A: Accession: A01763
A: Molecule type: protein
A: Residues: 1-26 <KRE>
C: Superfamily: mellitin major
C: Keywords: amidated carboxyl end; homotetramer
F: 26/Modified site: amidated carboxyl end (Glu) #status experimental

Query Match 61.7%; Score 87; DB 1; Length 26;
Best Local Similarity 76.0%; Pred. No. 7.1e-08;
Matches 19; Conservative 2; Mismatches 0; Indels 4; Gaps 1;

Oy 1 GIGAVLXXXXXVLTGTGALPALSMTK 25
Db 1 GIGAVLK----VLTGTGALPALSMTK 21
||||| |||||||||

RESULT 5
MEHBCF
mellitin - little honeybee
C: Species: Apis mellifera florea (little honeybee)
C: Date: 22-Jun-1981 #sequence_revision 22-Jun-1981 #text_change 15-Oct-1996
C: Accession: A01765
R: Kreil, G.
FEBS Lett. 33, 241-244, 1973
A: Title: Structure of mellitin isolated from two species of honey bees.
A: Reference number: A91392
A: Accession: A01765
A: Molecule type: protein
A: Residues: 1-26 <KRE>
C: Superfamily: mellitin major
C: Keywords: amidated carboxyl end; hemolysis; homotetramer; venom
F: 26/Modified site: amidated carboxyl end (Gln) #status experimental

Query Match 58.2%; Score 82; DB 1; Length 26;
Best Local Similarity 72.0%; Pred. No. 4.4e-07;
Matches 18; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

```

OY 1 GIGAVLKXXXVLTGGLPALISMIK 25
 |||||:|||||
 DB 1 GIGAILK-----VLTATGLPTLISMIK 21

RESULT 6

16534 renal sodium-dependent phosphate transporter type II - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 05-Nov-1999
 C:Accession: 146534
 R:Verri, T., Markovich, D., Perego, C., Norbis, F., Stange, G., Sorribas, V., Biber, J., Am, J. Physiol. 268, F626-F633, 1995
 A:Title: Cloning of a rabbit renal Na-Pi cotransporter, which is regulated by dietary pH
 A:Reference number: 146534; MUID:95251013; PMID:7733319
 A:Accession: 146534
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-642 <VER>
 A:Cross-references: EMBL:U20793; NID:9687736; PIDN:AAAY7682.1; PID:9687737

Query Match 32.6% Score 46; DB 2; Length 642;
 Best Local Similarity 40.9%; Pred. No. 14;
 Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 3 GAVLKXXXVLTGGLPALISMI 24
 |||||:|||||
 DB 382 GQVAVIKQVINTDLPAFTWV 403

RESULT 7

A10115 conserved hypothetical protein YP00945 [imported] - Yersinia pestis (strain CO92)
 C:Species: Yersinia pestis
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001
 C:Accession: A10115
 R:Parkhill, J., Wren, B.W., Thomson, N.R., Titball, R.W., Holden, M.T.G., Pentice, M.B., deno-Tarraga, A.M., Chillingworth, T., Cronin, A., Davies, R.M., Davis, P., Dougan, G., 11, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S., Barrett, Nature 413, 523-527, 2001
 A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
 A:Reference number: AB0001; MUID:21470413; PMID:11586360
 A:Accession: A10115
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-197 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC89788.1; PID:q15979015; GSPDB:GN00175
 C:Genetics:
 A:Gene: YP00945
 C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0226

Query Match 31.9% Score 45; DB 2; Length 197;
 Best Local Similarity 52.9%; Pred. No. 5.9;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 4 AVLKXXXVLTGGLPAL 20
 |||||:|||||
 DB 50 AILKARHMAQTGLPAI 66

RESULT 8

S11137 class I histocompatibility antigen alpha chain - cotton-top tamarin
 C:Species: Saguinus oedipus (cotton-top tamarin)
 C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 23-May-1997
 C:Accession: S11137
 R:Watkins, D.I., Chen, Z.W., Hughes, A.L., Evans, M.G., Tedder, T.F., Letvin, N.L., Nature 346, 60-63, 1990
 A:Title: Evolution of the MHC class I genes of a New World primate from ancestral homolo
 A:Reference number: S10934; MUID:9030971; PMID:2114550
 A:Accession: S11137
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tre
 A:Molecule type: mRNA

A:Residues: 1-357 <MAT>
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 F:212-277/Domain: immunoglobulin homology <IMM>

Query Match 31.9% Score 45; DB 2; Length 357;
 Best Local Similarity 29.7%; Pred. No. 11;
 Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

OY 1 GIGAVLKXXXVLTGGLPALISMIKXXXXXK 37
 |||||:|||||
 DB 302 GIVAILAILGAVVTGAVVAAMVRKSSDKGGSYSQ 338

RESULT 9

S11134 class I histocompatibility antigen alpha chain - cotton-top tamarin
 C:Species: Saguinus oedipus (cotton-top tamarin)
 C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 23-May-1997
 C:Accession: S11134
 R:Watkins, D.I., Chen, Z.W., Hughes, A.L., Evans, M.G., Tedder, T.F., Letvin, N.L., Nature 346, 60-63, 1990
 A:Title: Evolution of the MHC class I genes of a New World primate from ancestral hom
 A:Reference number: S10934; MUID:9030971; PMID:2114550
 A:Accession: S11134
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: mRNA
 A:Residues: 1-357 <MAT>
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 F:212-277/Domain: immunoglobulin homology <IMM>

Query Match 31.9% Score 45; DB 2; Length 357;
 Best Local Similarity 29.7%; Pred. No. 11;
 Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

OY 1 GIGAVLKXXXVLTGGLPALISMIKXXXXXK 37
 |||||:|||||
 DB 302 GIVAILAILGAVVTGAVVAAMVRKSSDKGGSYSQ 338

RESULT 10

S11135 class I histocompatibility antigen alpha chain - cotton-top tamarin
 C:Species: Saguinus oedipus (cotton-top tamarin)
 C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 23-May-1997
 C:Accession: S11135
 R:Watkins, D.I., Chen, Z.W., Hughes, A.L., Evans, M.G., Tedder, T.F., Letvin, N.L., Nature 346, 60-63, 1990
 A:Title: Evolution of the MHC class I genes of a New World primate from ancestral hom
 A:Reference number: S10934; MUID:9030971; PMID:2114550
 A:Accession: S11135
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
 A:Molecule type: mRNA
 A:Residues: 1-357 <MAT>
 C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
 F:212-277/Domain: immunoglobulin homology <IMM>

Query Match 31.9% Score 45; DB 2; Length 357;
 Best Local Similarity 29.7%; Pred. No. 11;
 Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

OY 1 GIGAVLKXXXVLTGGLPALISMIKXXXXXK 37
 |||||:|||||
 DB 302 GIVAILAILGAVVTGAVVAAMVRKSSDKGGSYSQ 338

RESULT 11

S11140 class I histocompatibility antigen alpha chain - cotton-top tamarin
 C:Species: Saguinus oedipus (cotton-top tamarin)
 C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 23-May-1997
 C:Accession: S11140
 R:Watkins, D.I., Chen, Z.W., Hughes, A.L., Evans, M.G., Tedder, T.F., Letvin, N.L., Nature 346, 60-63, 1990

A:Title: Evolution of the MHC class I genes of a New World primate from ancestral homology
A:Reference number: S10934; MUID:90305971; PMID:2114550
A:Accession: S11140
A:Status: preliminary
A:Molecule type: mRNA
A:Stucture: preliminary
A:Stucture: nucleic acid sequence not shown; not compared with conceptual trace
A:Residues: 1-357 <MAT>
A:Superfamily: class I histocompatibility antigen; immunoglobulin homology
F:212-277/Domatin: Immunoglobulin homology <IMM>

Query Match	31.98;	Score 45;	DB 2;	Length 357;
Best Local Similarity	29.78;	Pred. NO. 11;		
Matches 11;	Conservative 5;	Mismatches 21;	Indels 0;	Gaps 0;

```

Oy      1 GIGAVLKKXXXXVLTGTPALISWIKXXXXXXRRQ 37
         |||:|:|:|:|:|
Db      302 GIYAILALGAVVTGAVVAANVMRRKSSDKKGGSYQ 338

```

RESULT 12
S11139

class I histocompatibility antigen alpha chain - cotton-top tamarin
C:Species: *Saguinus oedipus* (cotton-top tamarin)
C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 23-May-1997
C:Accession: S11139
R:Wakibuchi, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.
Nature 346, 60-63, 1990
A:Title: Evolution of the MHC class I genes of a New World primate from ancestral hominoid
A:Reference number: S10934; MUID:90309971; PMID:2114550
A:Accession: S11139
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tree
A:Molecule type: mRNA
A:Residues: 1-357 <MAT>
C:Superfamily: class I histocompatibility antigen: Immunoglobulin homology
C:212-277/Domain: Immunoglobulin homology <IMM>

Query Match	31.9%	Score 45;	DB 2;	Length 357;
Best Local Similarity	29.7%	Pred. No. 11;		
Matches 11; Conservative	5;	Mismatches 21;	Indels 0;	Gaps 0

```
Oy 1 GIGAVLKKXXXXVLTGGLPALISWIKXXXXXXRRQ 37
    | | : | : | : | : |
Db 302 GIVALLAILGAVVTGAVVAAMVRRKSSDKKCGSYQ 338
```

RESULT 13
I72218

class I histocompatibility antigen I-C alpha chain precursor - cotton-top tamarin
N:Alternate names: MHC class IC protein
C:Species: *Saguinus oedipus* (cotton-top tamarin)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C:Accession: I72218: S11142
R:Wakkins, D.I.; Letvin, N.L.; Hughes, A.L.; Tedder, T.F.
J. Immunol. 144, 1136-1143, 1990
A:Title: Molecular cloning of cDNA that encode MHC class I molecules from a new world primate
A:Reference number: 156053; MUID:90111120; PMID:2104912
A:Accession: I72218
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-365 <RES>
A:Cross-references: GB:M3477; NID:g343319; PIDN:AAA36953.1; PID:g343320
R:Wakkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.
Nature 346, 60-63, 1990
A:Title: Evolution of the MHC class I genes of a New World primate from ancestral homol
A:Reference number: S10934; MUID:90309571; PMID:2114550
A:Accession: S11142
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 9-365 <MAT>
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: glycoprotein; transmembrane protein
F:220-285/Domain: Immunoglobulin homology <IMM>

Query match 31.98; Score 45; DB 2; Length 365;

	Best Local Similarity	29.7%	Pred. No. 11	
	Matches	11	Conservative	5
			Mismatches	21
			Indels	0
			Gaps	0
Oy	1	GIGAVLKKXXXVLTGTPALISWIKXXXXXXRQ	37	
		:	: :	: :
Db	310	GIVALLALIGAVVTGAVVAAMVMMRRKSSLRKKGCSISQ	346	

RESULT 14
S11133

class I histocompatibility antigen alpha chain - cotton-top tamarin
 C:Species: *Seiurus oedipus* (cotton-top tamarin)
 C:Date: 25-Feb-1994 #sequence_revision 26-May-1995 #text_change 23-May-1997
 C:Accession: S11133
 C:Wakkins, D.I.; Chem, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Lelvin, N.L.
 Nature 346, 60-63, 1990
 A:Title: Evolution of the MHC class I genes of a New World primate from ancestral hom
 A:Reference number: S10934; MUID:90306971; PMID:2114550

A; Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

Accession: 529 2811
 C: Superfamily: class I histocompatibility antigen: immunoglobulin homology
 F: 212-277/Domain: immunoglobulin homology <IMM>

Query Match	31.28;	Score 44;	DB 2;	Length 357;
Best Local Similarity	29.78;	Pred. No. 16;		
Matches	11;	Conservative	5;	Mismatches 21;
				Indels 0;
				Gaps 0

QY 1 GIGAVLKKXXXVLTGGLPALISWIKXXXXXXXXXRRQ 37
|||:| :| :| :| :|
Db 302 GIIVAILAIGVVVTGAVVAAMVRKSSDKKGGSYSO 338

RESULT 15
156053

class I histocompatibility antigen I-A alpha chain precursor - cotton-top tamarin
N:Alternate names: MHC class IA protein
C:Species: *Saguinus oedipus* (cotton-top tamarin)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Jul-1999
C:Accession: 156053; S10934
R:Watkins, D.I.; Letvin, N.L.; Hughes, A.L.; Tedder, T.F.
J. Immunol. 144, 1136-1143, 1990
A:Title: Molecular cloning of cDNA that encode MHC class I molecules from a new world
A:Reference number: 156053; MUID:90111120; PMID:2104912
A:Accession: 156053
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-365 <RES>
A:Cross-references: GB:M33475; NID:g343315; PIDN:AAA36951.1; PID:g343316
R:Watkins, D.I.; Chen, Z.W.; Hughes, A.L.; Evans, M.G.; Tedder, T.F.; Letvin, N.L.
Nature 346, 60-63, 1990
A:Title: Evolution of the MHC class I genes of a New World primate from ancestral hom
A:Reference number: S10934; MUID:90305971; PMID:2114550
A:Accession: S10934
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A:Molecule type: mRNA
A:Residues: 9-365 <MAT>
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
C:Keywords: glycoprotein; transmembrane protein
F:220-285/Domain: immunoglobulin homology <IMM>

Query Match	31.28;	Score 44;	DB 2;	Length 365;
Best Local Similarity	29.78;	Pred. No. 16;		
Matches	11;	Conservative	5;	Mismatches 21;
			Indels	0;
			Gaps	0

```
OY      1 GIGAVLKKXXXXXVLTGLPALISWIKXXXXXXXXXXRQ 37
        |||:|| : | : | : |
Db     310 GIALLAILGVVVGTGAVVAAMVMRRKSSDKKGGCSYSQ 346
```

Search completed: May 9, 2003, 16:18:46
Job time : 25.8431 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:08:20 ; Search time 12.2941 Seconds

(without alignments)
128.200 Million cell updates/sec

Title: US-09-851-422a-2

Perfect score: 141

Sequence: 1 GIGAVLKKXXXVLTGLPALISWIKXXXRXXXXXXRQ 38

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	66.0	70	1	MEL_APIME
2	87	61.7	26	1	MEL_APIDO
3	82	58.2	26	1	MEL_APIFL
4	51	36.2	22	1	MEL_RANTE
5	46	32.6	642	1	NPT2_RABIT
6	45	31.9	365	1	ICOL_SAGOE
7	44	31.2	365	1	LAOL_SAGOE
8	44	31.2	4344	1	DYHC_EMENT
9	43	30.5	303	1	SFR1_ARATH
10	43	30.5	445	1	DHFM_ECOLI
11	43	30.5	504	1	DEGP_RHIME
12	42.5	30.1	1081	1	CARB_SYNY3
13	42	29.8	639	1	NPT2_HUMAN
14	41	29.1	346	1	LUXA_PHOPO
15	41	29.1	354	1	LXAL_PHOPE
16	41	29.1	362	1	HLAH_HUMAN
17	41	29.1	365	1	LBOL_SAGOE
18	41	29.1	376	1	CYB_PLACH
19	41	29.1	551	1	OCRB_MYCUE
20	41	29.1	3951	1	VGF1_LBVB
21	40.5	28.7	393	1	YHFC_ECOLI
22	40.5	28.7	651	1	PTSA_PEDPE
23	40.5	28.7	761	1	NEMR_MOUSE
24	40	28.4	355	1	LUXA_VIBHA
25	40	28.4	360	1	LXAL_PHOLO
26	40	28.4	362	1	LXAL_PHOLO
27	40	28.4	509	1	DHNA_BACSU
28	40	28.4	519	1	DHNA_BACSP
29	40	28.4	543	1	UBIB_YERPE
30	40	28.4	546	1	UBIB_ECOLI
31	40	28.4	546	1	UBIB_SALTY
32	40	28.4	623	1	FAT1_YEAST
33	39.5	28.0	416	1	TFXD_RHILF

34	39	27.7	209	1	BACR_HALHS
35	39	27.7	217	1	NUZM_MYTED
36	39	27.7	258	1	TAM_DEIRA
37	39	27.7	313	1	NUDC_MYCTU
38	39	27.7	354	1	LUXA_VIBFI
39	39	27.7	357	1	LUXA_KRYAS
40	39	27.7	362	1	LOKO_GORCO
41	39	27.7	383	1	G22C_DROME
42	39	27.7	489	1	VGIC_MOPET
43	39	27.7	549	1	OCRB_MYCUE
44	39	27.7	836	1	VG26_BPMUS
45	39	27.7	2035	1	HFCL_HUMAN

ALIGNMENTS

RESULT 1	MEL_APIME	STANDARD:	PRT:	70 AA.
AC	P01501: P01503;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	21-JUL-1986 (Rel. 01, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Melittin precursor (Allergen Api m 3) (Api m III).			
OS	Apis mellifera (Honeybee), and			
OS	Apis cerana (Indian honeybee).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			
OC	Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;			
OC	Aculeata; Apoidea; Apidae; Apis.			
OX	NCBI_TaxID=7460, 7461;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES=A.mellifera;			
RX	MEDLINE=83287387; PubMed=6309516;			
RA	Vlasak R., Unger-Ullmann C., Kreil G., Frischaut A.-M.;			
RT	*Nucleotide sequence of cloned cDNA coding for honeybee			
RT	prepro-melittin.";			
RL	Eur. J. Biochem. 135:123-126(1983).			
RN	[2]			
RP	SEQUENCE OF 44-69 (MELITTIN I AND 2).			
RC	SPECIES=A.mellifera;			
RX	MEDLINE=68327913; PubMed=5592400;			
RA	Habermann E., Jentsch J.;			
RT	*Sequence analysis of melittin from tryptic and peptic degradation			
RT	products.";			
RL	Hoppe-Seyler's Z. Physiol. Chem. 348:37-50(1967).			
RN	[3]			
RP	SYNTHESIS.			
RC	SPECIES=A.mellifera;			
RX	MEDLINE=72098668; PubMed=5139482;			
RA	Schroeder E., Luebke K., Lehmann M., Beetz T.;			
RT	*Haemolytic activity and action on the surface tension of aqueous			
RT	solutions of synthetic melittins and their derivatives.";			
RL	Experientia 27:764-765(1971).			
RN	[4]			
RP	SYNTHESIS.			
RC	SPECIES=A.mellifera;			
RX	MEDLINE=72098669; PubMed=5139483;			
RA	Luebke K., Matthes S., Kloss G.;			
RT	*Isolation and structure of N 1-formyl melittin.";			
RL	Experientia 27:765-767(1971).			
RN	[5]			
RP	SEQUENCE OF 44-69.			
RC	SPECIES=A.cerana;			
RA	Krell G.;			
RT	*Structure of melittin isolated from two species of honey bees.";			
RL	FEBS Lett. 33:241-244(1973).			
RN	[6]			
RP	X-RAY CRYSTALLOGRAPHY.			
RC	SPECIES=A.mellifera;			
RX	MEDLINE=82189959; PubMed=7076662;			
RA	Terwilliger T.C., Eisenberg D.;			

```

RT "The structure of mellitin. II. Interpretation of the structure."
RL J. Biol. Chem. 257:6016-6022(1982).
RN [7]
RP STRUCTURE BY NMR OF 44-69.
RC SPECIES-A.mellifera;
RA Barham K.J., Hewish D., Wermelster J., Curtin C., Kirkpatrick A.,
RL Submitted (JUN-1998) to the PDB data bank.
RM [8]
RX MEDLINE=90254148; PubMed=2187536;
RA Dempsey C.E.;
RT "The actions of mellitin on membranes."
RL Biochim. Biophys. Acta 1031:143-161(1990).
CC -1- FUNCTION: MAIN TOXIN OF BEE VENOM WITH STRONG HEMOLYTIC ACTIVITY.
CC INTEGRATES INTO CELL MEMBRANES AND HAS MULTIPLE EFFECTS, PROBABLY,
CC AS A RESULT OF ITS INTERACTION WITH NEGATIVELY CHARGED
CC PHOSPHOLIPIDS. IT INHIBITS WELL KNOWN TRANSPORT PUMPS SUCH AS THE
CC NA(+)-K(+)-ATPASE AND THE H(+)-K(+)-ATPASE. INCREASES THE
CC PERMEABILITY OF CELL MEMBRANES TO IONS, PARTICULARLY NA+ AND
CC INDIRECTLY CA2+, BECAUSE OF THE NA(+)-CA(2+)-EXCHANGE.
CC -1- SUBUNIT: MONOMER AND HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- MISCELLANEOUS: N-FORMYL-MELLITIN MAJOR HAS 80% OF THE ACTIVITY OF
CC MELLITIN.
CC -1- DATABASE: NAME-Protein Spotlight;
CC NOTE-Issue 12 of July 2001;
CC WWW="http://www.expasy.org/spotlight/articles/spl1012.html".
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X02007; CAA26038.1;
DR PIR: A01761; MPB1.
DR PIR: A01764; MEHB2.
DR PIR: A01762; MEHB1.
DR PDB: 2MLT; 15-JUL-92.
DR PDB: 1BH1; 06-JAN-99.
DR InterPro: IPR002116; Mellitin.
DR Pfam: PF01372; Mellitin; 1.
DR ProDom: PD014636; Mellitin; 1.
DR Hemolysis: Toxin; Amidation; Formylation; Signal; 3D-structure;
KW Allergen.
FT SIGNAL 1 21
FT PROPEP 22 43 REMOVED BY A DIPEPTIDYLPEPTIDASE.
FT CHAIN 44 69 MELLITIN.
FT MOD_RES 44 44 FORMYLATION (ABOUT 10% OF THE MOLECULES).
FT MOD_RES 69 69 AMIDATION (G-70 PROVIDE AMIDE GROUP).
FT VARIANT 64 64 K -> S (IN MELLITIN 2; POSSIBLY AN
FT ARTIFACT).
FT VARIANT 67 70 KQOG -> KRQO (IN MELLITIN 2; POSSIBLY AN
FT ARTIFACT).
FT HELIX 45 53
FT TURN 54 54
FT HELIX 55 58
SQ SEQUENCE 70 AA; 7585 MW; 607F52C091C3BB6 CRC64;
Query Match 66.0%; Score 93; DB 1; Length 70;
Best Local Similarity 68.4%; Pred. No. 6.4e-09;
Matches 26; Conservative 0; Mismatches 0; Indels 12; Gaps 3;
OY 1 GIGAVLXXXXVLTGTLPLISWIKXXXXXKROQ 38
DB 44 GIGAVL-----VLTGTLPLISWIK-----RK----RQO 69
RESULT 2
MEL_APIDO

```

```

ID MEL_APIDO STANDARD: PRT; 26 AA.
AC P01502;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mellitin.
OS Apis dorsata (Giant honeybee).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC Aculeata; Apoidea; Apidae; Apis.
OX NCBI_TaxID=7463;
RN [1]
RP SEQUENCE.
RX MEDLINE=75168194; PubMed=1093875;
RA Krell G.;
RT "The structure of Apis dorsata mellitin: phylogenetic relationships
RL between honeybees as deduced from sequence data."
RL FEBS Lett. 54:100-102(1975).
CC -1- FUNCTION: MAIN TOXIN OF BEE VENOM WITH STRONG HEMOLYTIC ACTIVITY.
CC INTEGRATES INTO CELL MEMBRANES AND HAS MULTIPLE EFFECTS, PROBABLY,
CC AS A RESULT OF ITS INTERACTION WITH NEGATIVELY CHARGED
CC PHOSPHOLIPIDS. IT INHIBITS WELL KNOWN TRANSPORT PUMPS SUCH AS THE
CC NA(+)-K(+)-ATPASE AND THE H(+)-K(+)-ATPASE. INCREASES THE
CC PERMEABILITY OF CELL MEMBRANES TO IONS, PARTICULARLY NA+ AND
CC INDIRECTLY CA2+, BECAUSE OF THE NA(+)-CA(2+)-EXCHANGE.
CC -1- SUBUNIT: MONOMER AND HOMOTETRAMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DATABASE: NAME-Protein Spotlight;
CC NOTE-Issue 12 of July 2001;
CC WWW="http://www.expasy.org/spotlight/articles/spl1012.html".
CC -----
DR PIR: A01763; MEHCD.
DR HSSP: P01501; 1BH1.
DR InterPro: IPR002116; Mellitin.
DR Pfam: PF01372; Mellitin; 1.
DR ProDom: PD014636; Mellitin; 1.
DR Hemolysis: Toxin; Amidation.
FT MOD_RES 26 26
FT SEQUENCE 26 AA; 2848 MW; F1DA8F92514EF01C CRC64;
SQ SEQUENCE 26 AA; 7585 MW; 607F52C091C3BB6 CRC64;
Query Match 61.7%; Score 87; DB 1; Length 26;
Best Local Similarity 76.0%; Pred. No. 2.5e-08;
Matches 19; Conservative 2; Mismatches 0; Indels 4; Gaps 1;
OY 1 GIGAVLXXXXVLTGTLPLISWIK 25
DB 1 GIGAVL-----VLTGTLPLISWIK 21
RESULT 3
MEL_APIFL
ID MEL_APIFL STANDARD: PRT; 26 AA.
AC P01504;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mellitin.
OS Apis florea (Little honeybee).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita;
OC Aculeata; Apoidea; Apidae; Apis.
OX NCBI_TaxID=7463;
RN [1]
RP SEQUENCE.
RA Krell G.;
RT "Structure of mellitin isolated from two species of honey bees."
RL FEBS Lett. 33:241-244(1973).
CC -1- FUNCTION: MAIN TOXIN OF BEE VENOM WITH STRONG HEMOLYTIC ACTIVITY.
CC INTEGRATES INTO CELL MEMBRANES AND HAS MULTIPLE EFFECTS, PROBABLY,
CC AS A RESULT OF ITS INTERACTION WITH NEGATIVELY CHARGED
CC PHOSPHOLIPIDS. IT INHIBITS WELL KNOWN TRANSPORT PUMPS SUCH AS THE
CC NA(+)-K(+)-ATPASE AND THE H(+)-K(+)-ATPASE. INCREASES THE
CC PERMEABILITY OF CELL MEMBRANES TO IONS, PARTICULARLY NA+ AND
CC INDIRECTLY CA2+, BECAUSE OF THE NA(+)-CA(2+)-EXCHANGE.

```

CC -1- SUBUNIT: MONOMER AND HOMOTETRAMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DATABASE: NAME-Protein Spotlight;
CC NOTE-Issue 12 of July 2001:
CC WWW="http://www.expasy.org/spotlight/articles/sp1c1012.html".
DR PIR: A01765; MEMBCF.
DR HSSP: P01501; 1BHI.
DR InterPro: IPR002116; Melittin.
DR Pfam: PF01372; Melittin; 1.
DR ProDom: PD014636; Melittin; 1.
KW Hemolysis; Toxin; Amidation.
FT MOD_RES 26
SQ SEQUENCE 26 AA: 2819 MW: F20E8F82400EF01C CRC64;
Query Match 58.2% Score 82; DB 1; Length 26;
Best Local Similarity 72.0% Pred. No. 1.7e-07;
Matches 18; Conservative 1; Mismatches 2; Indels 4; Gaps 1;
QY 1 GIGAVLKKXXVLTGTPALISWIK 25
Db 1 GIGAVLKKXXVLTGTPALISWIK 21
RESULT 4
MLP_RANTE STANDARD: PRT: 22 AA.
ID MLP_RANTE
AC P56924;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melittin-like peptide (MLP).
OS Rana temporaria (European common frog).
CC Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8407;
RN [1]
RP SEQUENCE.
RC TISSUE=SKIN;
RX MEDLINE=97175050; PubMed=9022710;
RA Simmaco M., Mignogna G., Canofeni S., Miele R., Mangoni M.L.,
RT Barra D.;
RT "temporaria, antimicrobial peptides from the European red frog Rana
temporaria.";
RL Eur. J. Biochem. 242:788-792(1996).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: HIGH, TO A.MILLIFERA MELITTIN.
KW Amphibian skin; Amidation.
FT MOD_RES 22
SQ SEQUENCE 22 AA: 2313 MW: 0990320F9B0709B3 CRC64;
Query Match 36.2% Score 51; DB 1; Length 22;
Best Local Similarity 45.8% Pred. No. 0.025;
Matches 11; Conservative 5; Mismatches 4; Indels 4; Gaps 1;
QY 2 IGAVLKKXXVLTGTPALISWIK 25
Db 2 IGAVLKKXXVLTGTPALISWIK 21
RESULT 5
NPT2_RABIT STANDARD: PRT: 642 AA.
ID NPT2_RABIT
AC Q28620;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Renal sodium-dependent phosphate transporter 2 (Sodium/phosphate
cotransporter 2) (Na(+)/Pi cotransporter 2) (Renal sodium-phosphate
transport protein 2) (Renal Na(+)-dependent phosphate cotransporter 2)
(NaPi-6).
GN SLC34A1 OR SLC17A2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Euthera; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=95251013; PubMed=7733319;
RA Verri T., Markovich D., Perego C., Norbis F., Stange G., Sorribas V.,
RA Biber J., Murer H.;
RT "Cloning of a rabbit renal Na-Pi cotransporter, which is regulated by
dietary phosphate.";
RL Am. J. Physiol. 268:F626-F633(1995).
CC -1- FUNCTION: MAY BE INVOLVED IN ACTIVE TRANSPORTING PHOSPHATE
INTO CELLS VIA NA+ COTRANSPORT IN THE RENAL BRUSH BORDER
MEMBRANE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL: U20793; AAA77682.1;
DR InterPro: IPR003841; Na_Pi_cotrans.
DR Pfam: PF02690; Na_Pi_cotrans; 1.
DR TIGRFAMs: TIGR01013; 2858; 1.
KW Transport; Symport; Sodium transport; Transmembrane; Glycoprotein;
KW Phosphorylation.
FT DOMAIN 1 106
FT TRANSMEM 107 128
FT DOMAIN 129 148
FT TRANSMEM 149 166
FT DOMAIN 167 168
FT TRANSMEM 169 188
FT DOMAIN 189 350
FT TRANSMEM 351 373
FT DOMAIN 374 415
FT TRANSMEM 416 439
FT DOMAIN 440 469
FT TRANSMEM 470 490
FT DOMAIN 491 516
FT TRANSMEM 517 537
FT DOMAIN 538 542
FT TRANSMEM 543 564
FT DOMAIN 565 642
FT DOMAIN 643 672
FT MOD_RES 511 511
FT CARBOHYD 301 301
FT CARBOHYD 326 326
FT CARBOHYD 333 333
SQ SEQUENCE 642 AA: 69431 MW: A48B85C001642A8B CRC64;
Query Match 32.6% Score 46; DB 1; Length 642;
Best Local Similarity 40.9% Pred. No. 4.9;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
QY 3 GAVLKKXXVLTGTPALISWIK 24
Db 382 GOVAKVIOKVIWTDLPAPFTW 403
RESULT 6
ICOL_SAGE STANDARD: PRT: 365 AA.
ID ICOL_SAGE
AC P30517;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-APR-1993 (Rel. 25, Last annotation update)
DE Class I histocompatibility antigen, C alpha chain precursor.
OS Saguinus oedipus (Cotton-top tamarin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

ORGANELLES ALONG MICROTUBULES. REQUIRED TO MAINTAIN UNIFORM
NUCLEAR DISTRIBUTION IN HYPHAE.
-1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
INTERMEDIATE AND LIGHT CHAINS.
-1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U03904; AAA18338.1; -
DR PIR: A53489; A53489.
DR InterPro: IPR004273; Dynein_heavy.
KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.
FT DOMAIN 667 687 COILED COIL (POTENTIAL).
FT DOMAIN 913 933 COILED COIL (POTENTIAL).
FT DOMAIN 1321 1341 COILED COIL (POTENTIAL).
FT DOMAIN 1548 1564 COILED COIL (POTENTIAL).
FT DOMAIN 1627 1656 COILED COIL (POTENTIAL).
FT DOMAIN 2035 2063 MICROTUBULE-BINDING (POTENTIAL).
FT DOMAIN 3179 3270 COILED COIL (POTENTIAL).
FT DOMAIN 3407 3463 COILED COIL (POTENTIAL).
FT DOMAIN 3700 3798 COILED COIL (POTENTIAL).
FT NP_BIND 1933 1940 ATP (POTENTIAL).
FT NP_BIND 2223 2230 ATP (POTENTIAL).
FT NP_BIND 2592 2599 ATP (POTENTIAL).
FT NP_BIND 2932 2939 ATP (POTENTIAL).
SQ SEQUENCE 4344 AA: 492470 MW: 30245864cF8BD019 CRC64;
Query Match 31.2% Score 44; DB 1; Length 4344;
Best Local Similarity 41.2% Pred. No. 71;
Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
OY 15 TGLPALISWIKKXXRK 31
DB 4284 TGLPAFVEMWNLPERE 4300

RESULT 9
SPRL ARATH STANDARD: PRT: 303 AA.
ID SFRI_ARATH 022315: Q22314; Q39201;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pre-mRNA splicing factor SF2 (SRI protein).
GN SF2 OR AT1G02840 OR F22D16.16 OR F22D16_30.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_Taxid=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95372342; PubMed-7644475;
RA Lazar G., Schaal T., Maniatis T., Goodman H.M.;
RT "Identification of a plant serine-arginine-rich protein similar to the
RT mammalian splicing factor SF2/ASF".
RL Proc. Natl. Acad. Sci. U.S.A. 92:7672-7676(1995).
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Su C.-L., Schuler M.A.;
RL Submitted (Apr-1997) to the EMBL/Genbank/DDB databases.
RN [13]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;

MEDLINE-21016719; PubMed-11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altieri H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Feng B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Malt R., Marzilli A.,
RA Maltischer J., Miranda M., Nguyen M., Niernan W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
RA Unterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Frazer C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana".
RL Nature 408:816-820(2000).
CC -1- FUNCTION: CAN PROMOTE SPICE SITE SELECTION IN VITRO PRESUMABLY BY
CC ANTICONTINIZING THE EFFECTS OF THE AT HETEROGENEOUS NUCLEAR
CC RIBONUCLOPROTEIN. MAY HAVE AN ESSENTIAL FUNCTION DURING EARLY
CC PLANT DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM/SF2A (SHOWN HERE)
CC AND A SHORT FORM/SF2B: ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: UBIVITOUS.
CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS
CC DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M98340; AAA32856.1; -
DR EMBL: AF001035; AAB71385.1; -
DR EMBL: AF001035; AAB71386.1; -
DR EMBL: AC009525; AAF02881.1; -
DR HSSP: P19339; ISXL.
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00360; RRM; 2.
DR PROSITE: PS00360; RRM; 2.
DR PROSITE: PS00360; RRM; 2.
KW mRNA processing; mRNA splicing; Nuclear protein; RNA-binding; Repeat;
KW Phosphorylation; Alternative splicing.
FT DOMAIN 7 82 RNA-BINDING (RRM) 1.
FT DOMAIN 96 112 GLY-RICH (HINGE REGION).
FT DOMAIN 119 197 RNA-BINDING (RRM) 2.
FT DOMAIN 198 273 ARG/SER-RICH (RS DOMAIN).
FT DOMAIN 274 303 LYS/PRO/SER-RICH (PSK DOMAIN).
FT VARSPLIC 267 272 SRSRSR -> YGFTYD (IN ISOFORM SF2B).
FT VARSPLIC 273 303 MISSING (IN ISOFORM SF2B).
FT CONFLICT 125 125 G -> W (IN REF. 1).
FT CONFLICT 127 127 P -> A (IN REF. 1).
FT CONFLICT 139 140 MR -> IA (IN REF. 1).
SQ SEQUENCE 303 AA: 33729 MW: 6530F9CB628BBEFA CRC64;
Query Match 30.5% Score 43; DB 1; Length 303;
Best Local Similarity 66.7% Pred. No. 7.5;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 12 VLTGTPALISW 23
DB 121 VLTGTPSSASW 132

```

RESULT 10
YHFM_ECOLI STANDARD: PRT: 445 AA.
ID YHFM_ECOLI
DE P45539: P76686;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical transport protein yhfM.
CN YHFM OR B3370.
OS Escherichia coli.
OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae:
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RL "The complete genome sequence of Escherichia coli K-12."
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
CC PERMEASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
-----
DR EMBL: U18997; AA58167.1; ALT_SEQ.
DR EMBL: AE000413; AAC76395.1; ALT_INIT.
DR EcoGene: EG12908; yhfM.
DR InterPro: IPR002293; AA/rel_prtmase1.
DR InterPro: IPR004841; Permease.
DR Pfam: PF00324; aa_permeases.1.
KW Hypothetical protein; transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 10 30 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 121 141 POTENTIAL.
FT TRANSMEM 155 175 POTENTIAL.
FT TRANSMEM 181 201 POTENTIAL.
FT TRANSMEM 236 256 POTENTIAL.
FT TRANSMEM 273 293 POTENTIAL.
FT TRANSMEM 334 354 POTENTIAL.
FT TRANSMEM 355 375 POTENTIAL.
FT TRANSMEM 389 410 POTENTIAL.
FT TRANSMEM 417 435 POTENTIAL.
FT TRANSMEM 445 AA; 47576 MW; B98DF5944A5E1FEF CRC64;
SQ SEQUENCE

```

```

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable serine protease do-1ike precursor (EC 3.4.21.-).
CN DEGP1 OR DEGP OR R01021 OR SMC02365.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria: Proteobacteria: alpha subdivision: Rhizobiaceae group:
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=96146524; PubMed=8550509;
RA Glazebrook J., Ichige A., Walker G.C.;
RT "Genetic analysis of Rhizobium meliloti bacA-phoA fusion results in
RT identification of degp: two loci required for symbiosis are closely
RT linked to degp."
RL J. Bacteriol. 178:745-752(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Cordie T., Coffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puehler A., Fumelle B., Rampeger U.,
RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -1- SUBCELLULAR LOCATION: Periplasmic (potential).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2C.
CC -1- SIMILARITY: CONTAINS 2 PDZ/DR DOMAINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch)
-----
DR EMBL: U31512; AAC43669.1; ALT_INIT.
DR EMBL: AL591785; CAC45593.1; -.
DR MEROPS: S01.273; -.
DR InterPro: IPR001478; PDZ.
DR InterPro: IPR001940; Protease2C.
DR InterPro: IPR000126; Ser-protease_V8.
DR InterPro: IPR001254; Ser-protease_Try.
DR Pfam: PF00089; trypsin.1.
DR Pfam: PF00595; PDZ_2.
DR PRINTS: PR00834; PROTEASES2C.
DR PRINTS: PR00839; V8PROTEASE.
DR SMART: SM00228; PDZ_2.
DR PROSITE: PS50106; PDZ_2.
KW Hydrolase; Serine protease; Periplasmic; Repeat; Signal;
KW Complete proteome.
FT SIGNAL 1 26
FT CHAIN 27 504
FT DOMAIN 113 286
FT DOMAIN 287 378
FT DOMAIN 401 491
FT ACT_SITE 140 140
FT ACT_SITE 170 170
FT ACT_SITE 244 244
FT ACT_SITE 244 244
FT ACT_SITE 14 15
FT CONFLICT 39 147
FT CONFLICT

```

RESULT 11
 DEGP_RHIME STANDARD: PRT: 504 AA.
 AC 052894:
 DT 30-MAY-2000 (rel. 39, Created)

Query Match
 Best Local Similarity 57.18; Score 43; DB 1; Length 445;
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
 Oy 12 VLTGGLPALISWIK 25
 I:|||||I
 Db 423 VIATGLPAVAFMAK 436

RESULT 11
 DEGP_RHIME STANDARD: PRT: 504 AA.
 AC 052894:
 DT 30-MAY-2000 (rel. 39, Created)

```

SQ SEQUENCE 504 AA: 53035 MW: D7E82B89981EA23C CRC64:
Query Match 30.5% Score 43: DB 1; Length 504:
Best Local Similarity 50.0% Pred. No. 12:
Matches 11: Conservative 3: Mismatches 2: Indels 6: Gaps 1:
OY 1 GIGAVLKKXXXXXVLTGTPALIS 22
   1: ||| |||||| ::
DB 9 GLAVAL-----LTTGLPAEVA 24

RESULT 12
CARB_SYNY3 STANDARD: PRT: 1081 AA.
AC Q55736:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-
DE phosphate synthetase ammonia chain).
GN CARB OR SLL0370.
OS Synchocystis sp. (Strain PCC 6803):
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Tabeiuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synchocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions." ;
RL DNA Res. 3:109-136(1996).
CC -1- CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -1- COFACTOR: Binds three manganese ions (By similarity).
CC -1- PATHWAY: Arginine biosynthesis.
CC -1- PATHWAY: Pyrimidine biosynthesis; first step.
CC -1- SUBUNIT: Composed of two chains; the small (or glutamine) chain
CC promotes the hydrolysis of glutamine to ammonia, which is used by
CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By
CC similarity).
CC -1- SIMILARITY: BELONGS TO THE CARB FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D64002; BAA10403.1; ALT_INITY.
DR HSSP: P00968; IAGX.
DR InterPro: IPR005483; CPase_L.
DR InterPro: IPR005479; CPase_L_D2.
DR InterPro: IPR005480; CPase_L_D3.
DR InterPro: IPR005481; CPase_L_N.
DR InterPro: IPR004362; MGS_Like.
DR Pfam: PF00289; CPase_L_chain; 2.
DR Pfam: PF02786; CPase_L_D2; 2.
DR Pfam: PF02787; CPase_L_D3; 1.
DR Pfam: PF02142; MGS; 1.
DR PRINTS: PR00098; CPASE.
DR PROSITE: PS00866; CPASE_1; 1.
DR PROSITE: PS00867; CPASE_2; 2.
KW Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;
KW ATP-binding; Manganese; Complete proteome.
FT DOMAIN 1 403
FT DOMAIN 404 553
FT OLIGOMERIZATION DOMAIN.

```

[illegible]

FT DOMAIN 164 165 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 166 185 M3 (POTENTIAL).
 FT DOMAIN 186 347 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 348 370 M4 (POTENTIAL).
 FT DOMAIN 371 412 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 413 436 M5 (POTENTIAL).
 FT DOMAIN 437 466 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 467 487 M6 (POTENTIAL).
 FT DOMAIN 488 513 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 514 534 M7 (POTENTIAL).
 FT DOMAIN 535 539 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 540 561 M8 (POTENTIAL).
 FT DOMAIN 562 639 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 639 652 LEUCINE-ZIPPER.
 FT MOD_RES 508 508 PHOSPHORYLATION (BY PKC) (POTENTIAL).
 FT CARBOHYD 298 298 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT SEQUENCE 639 AA: 68937 MW: 65D21D968C35D61B CRC64;

Query Match 29.8%; Score 42; DB 1; Length 639;
 Best Local Similarity 36.4%; Pred. No. 23;
 Matches 8; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 3 GAVLKXXXXVLTGLPALISW 24
 DB 379 GOVAVKVIQKINTDFPAPFTWV 400

RESULT 14
 ID LUXA_PHOPO STANDARD: PRT; 346 AA.
 AC P24113;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Alkanal monooxygenase alpha chain (EC 1.14.14.3) (Bacterial Luciferase
 DE alpha chain).
 GN LUXA.
 OS Photobacterium phosphoreum.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae;
 CC Photobacterium.
 CC NCBI_TaxID=659;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91207448; PubMed=2018544;
 RA Ferri S.R., Soly R.R., Szilner R.B., Meighen E.A.;
 RT "Structure and properties of luciferase from Photobacterium
 RT phosphoreum."
 RL Biochem. Biophys. Res. Commun. 176:541-548(1991).
 CC [2]
 RP SEQUENCE OF 1-41 FROM N.A.
 RX MEDLINE=91302295; PubMed=2071574;
 RA Ferri S.R., Meighen E.A.;
 RT "A lux-specific myristoyl transferase in luminescent bacteria related
 RT to eukaryotic serine esterases."
 RL J. Biol. Chem. 266:12852-12857(1991).
 CC [3]
 RP FUNCTION: LIGHT-EMITTING REACTION IN LUMINOUS BACTERIA.
 CC -1- CATALYTIC ACTIVITY: RCHO + FMNH(2) + O(2) -> RCOOH + FMN + H(2)O +
 CC light.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M65067; AAA70297.1; -;
 CC DR EMBL: M64224; AAA25627.1; -;
 CC DR PIR: JH0387; JH0387.

DR HSSP; P07740; LUUC.
 DR InterPro: IPR002103; Bac_luciferase.
 DR Pfam: PF00296; bac_luciferase; 1.
 DR PRINTS: PR00089; LUCIFERASE.
 DR PROSITE: PS00494; BACTERIAL_LUCIFERASE; 1.
 KW Photoprotein; Luminescence; Oxidoreductase; Monooxygenase;
 KW Flavoprotein; FMN.
 FT DOMAIN 97 111 REGION OF ACTIVE CENTER-1 (BY
 FT SIMILARITY).
 FT DOMAIN 273 288 REGION OF ACTIVE CENTER-2 (BY
 FT SIMILARITY).
 FT CONFLICT 6 6 I-> IC (IN REF. 2).
 FT CONFLICT 40 41 LE-> TD (IN REF. 2).
 FT SEQUENCE 346 AA: 39364 MW: 360EC324E0A438A0 CRC64;

Query Match 29.1%; Score 41; DB 1; Length 346;
 Best Local Similarity 43.8%; Pred. No. 19;
 Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 16 GLPALISWIKXXXXRK 31
 DB 180 GLPWLWLTITTSKK 195

RESULT 15
 ID LXA1_PHOLE STANDARD: PRT; 354 AA.
 AC P09140;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Alkanal monooxygenase alpha chain (EC 1.14.14.3) (Bacterial Luciferase
 DE alpha chain).
 GN LUXA.
 OS Photobacterium leiognathi.
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae;
 CC Photobacterium.
 CC NCBI_TaxID=658;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=554;
 RX MEDLINE=89041570; PubMed=3186447;
 RA Iltariinov B.A., Protodopova M.V., Karginov V.A., Mertvetsov N.P.,
 RA Gilels J.I.;
 RT "Nucleotide sequence of part of Photobacterium leiognathi lux
 RT region."
 RL Nucleic Acids Res. 16:9855-9855(1988).
 CC [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=554;
 RX MEDLINE=88251495; PubMed=3382442;
 RA Iltariinov B.A., Protodopova M.V., Karginov V.A., Mertvetsov N.P.,
 RA Gilels J.I.;
 RT "Nucleotide sequence of genes for alpha- and beta-subunits of
 RT luciferase from Photobacterium leiognathi."
 RL Bioorg. Khim. 14:412-415(1988).
 CC [3]
 RP SEQUENCE FROM N.A.
 RX STRAIN=554;
 RX MEDLINE=90185252; PubMed=2311938;
 RA Iltariinov B.A., Bilnov V.M., Donchenko A.P., Protodopova M.V.,
 RA Karginov V.A., Mertvetsov N.P., Gilels J.I.;
 RT "Isolation of bioluminescent functions from Photobacterium
 RT leiognathi: analysis of luxA, luxB, luxG and neighboring genes."
 RL Gene 86:89-94(1990).
 CC [4]
 RP FUNCTION: LIGHT-EMITTING REACTION IN LUMINOUS BACTERIA.
 CC -1- CATALYTIC ACTIVITY: RCHO + FMNH(2) + O(2) -> RCOOH + FMN + H(2)O +
 CC light.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X08036; CAA30831.1; -
DR PIR: S01697; S01697.
DR PIR: J00413; J00413.
DR PIR: S06369; S06369.
DR HSSP: P07740; 1BRL.
DR InterPro: IPR002103; Bac_luciferase.
DR Pfam: PF00296; Bac_luciferase; 1.
DR PRINTS: PR00089; LUCIFERASE.
DR PROSITE: PS00494; BACTERIAL_LUCIFERASE; 1.
KM Photoprotein: Luminescence; Oxidoreductase; Monooxygenase;
KW Flavoprotein; FMN.
FT DOMAIN 100 115 REGION OF ACTIVE CENTER-1 (BY
SIMILARITY).
FT DOMAIN 279 294 REGION OF ACTIVE CENTER-2 (BY
SIMILARITY).
SQ SEQUENCE 354 AA; 40370 MW; 569F6EEB903C7B90 CRC64;

Query Match 29.1% Score 41; DB 1; Length 354;
Best Local Similarity 43.8% Pred. No. 19;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 16 GLPALISWIKXXXXRK 31
|||::|||:
Db 186 GLPMVLSWITTSRK 201

Search completed: May 9, 2003, 16:16:18
Job time : 14.2941 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:10:40 ; Search time 26.0784 Seconds
(without alignments)
300.240 Million cell updates/sec

Title: US-09-851-422A-2
Perfect score: 141
Sequence: 1 GIGAVLKXXXVLTGTPALISWIKXXXXXXRQ 38

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._virus:*
16: sp._bacteria:*
17: sp._archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	66.0	67	5	Q95219
2	47	33.3	2159	10	Q8RV11
3	47	33.3	2159	10	Q8RV11
4	45	31.9	197	16	Q8ZHF4
5	45	31.9	320	16	Q8ZHF4
6	45	31.9	357	7	Q31016
7	45	31.9	357	7	Q31016
8	45	31.9	358	7	Q31012
9	44	31.2	101	11	Q9CTZ3
10	44	31.2	357	7	Q62903
11	44	31.2	457	17	Q8TNG7
12	44	31.2	714	5	Q18105
13	44	31.2	749	16	Q9RY43
14	44	31.2	762	16	Q9AT67
15	44	31.2	2078	16	Q98K31
16	43.5	30.9	200	16	Q9RXX6

17	43	30.5	174	16	Q97RS0	Q97RS0 streptococ
18	43	30.5	240	13	Q9DDR8	Q9DDR8 gallus gall
19	43	30.5	261	10	Q9SPI3	Q9SPI3 arabidopsis
20	43	30.5	270	10	Q9SPI5	Q9SPI5 arabidopsis
21	43	30.5	276	10	Q9SPI5	Q9SPI5 arabidopsis
22	43	30.5	283	16	Q9SP10	Q9SP10 mycobacteri
23	43	30.5	289	10	Q9SP12	Q9SP12 arabidopsis
24	43	30.5	307	10	Q9SP11	Q9SP11 arabidopsis
25	43	30.5	425	4	Q9UL02	Q9UL02 homo sapien
26	43	30.5	428	5	Q8SXA2	Q8SXA2 drosophila
27	43	30.5	440	2	Q93LM6	Q93LM6 nostoc punc
28	43	30.5	462	16	Q8X845	Q8X845 escherichia
29	43	30.5	658	5	Q9VN12	Q9VN12 drosophila
30	43	30.5	744	16	Q8Y4H8	Q8Y4H8 listeria mo
31	43	30.5	850	5	Q9Y137	Q9Y137 drosophila
32	42.5	30.1	330	16	Q49885	Q49885 mycobacteri
33	42.5	30.1	357	2	Q52098	Q52098 photobacter
34	42.5	30.1	359	2	Q93E16	Q93E16 rhizobium 1
35	42.5	30.1	1105	16	Q55756	Q55756 synechocyst
36	42	29.8	239	16	Q927J3	Q927J3 listeria in
37	42	29.8	250	10	Q8RZS6	Q8RZS6 oryza sativ
38	42	29.8	262	17	Q97U77	Q97U77 sulfolobus
39	42	29.8	425	11	Q9RYM8	Q9RYM8 mus musculu
40	42	29.8	425	11	Q64602	Q64602 r kynurenin
41	42	29.8	478	16	Q9H0Z2	Q9H0Z2 pseudomonas
42	42	29.8	667	5	Q93839	Q93839 caenorhabdi
43	42	29.8	1269	10	Q9FME9	Q9FME9 arabidopsis
44	42	29.8	9507	2	Q9EWA1	Q9EWA1 streptomyce
45	41	29.1	132	2	Q9K129	Q9K129 edwardsiell

ALIGNMENTS

RESULT 1	Q95219	PRELIMINARY:	PRT:	67 AA.
ID	Q95219			
AC	Q95219			
DT	01-DEC-2001 (TREMUR19, 19, Created)			
DT	01-DEC-2001 (TREMUR19, 19, Last sequence update)			
DT	01-JUN-2002 (TREMUR21, 21, Last annotation update)			
DE	Melittin protein.			
GN	MELITTIN.			
OS	Polistes sp. (Golden paper wasp).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;			
OC	Vespoidea; Vespidae; Polistinae; Polistes.			
OX	NCBI_TaxID=30210;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Le H.O.;			
RL	Thesis (2001), Department of Lab. Molecular Microbiology,			
RL	Institute of Biotechnology, Hanoi, Viet Nam.			
DR	EMBL; AJ314920; CAC42164.1;			
DR	InterPro: IPR002116; Melittin.			
DR	Pfam: PF01372; Melittin; 1.			
DR	ProDom: PD014636; Melittin; 1.			
FT	CHAIN 41 67			
FT	SEQUENCE 67 AA; 7502 MW; 94821CF0EAEE9CF CRC64;			
QY	1 GIGAVLKXXXVLTGTPALISWIKXXXXXXRQ 38			
Db	41 GIGAVLK-----VLTGTPALISWIK-----RK-----RQ 66			
Matches	26; Conservative			
Mismatches	0; Indels			
Gaps	12; Gaps			
Score	66.0%; Score 93; DB 5; Length 67;			
Best Local Similarity	68.4%; Pred. No. 4.2e-08;			
Result 2	Q8RV11	PRELIMINARY:	PRT:	2159 AA.
ID	Q8RV11			
AC	Q8RV11			

DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Calpain-like protein.
 GN DEK1.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoidae; Andropogoneae; Zea.
 OC NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. B73; TISSUE=DEVELOPING ENDOSPERM;
 RX PubMed=11929961;
 RA Lid S.E., Grulis D., Jung R., Lorentzen J.A., Ananiev E.,
 RA Chamberlin M., Niu X., Meeley R., Nichols S., Olsen O.A.;
 RT "The defective kernel 1 (dek1) gene required for aleurone cell
 RT development in the endosperm of maize grains encodes a membrane
 RT protein of the calpain gene superfamily."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:5460-5465(2002).
 DR EMBL: AY061806; AAL38189.1;
 SO SEQUENCE 2159 AA; 238997 MW; 9F995D3F9EAC1315 CRC64;

Query Match 33.3%; Score 47; DB 10; Length 2159;
 Best Local Similarity 36.4%; Pred. No. 79;
 Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 2 IGAUXXXXXXVLTGTPALISM 23
 Db 821 LGWALSTIAVITGTPALISM 842

RESULT 3
 ID 08RU01 PRELIMINARY; PRT: 2159 AA.

AC 08RU01:
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Calpain-like protein.
 GN DEK1.
 OS Zea mays (Maize).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoidae; Andropogoneae; Zea.
 OC NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. MO17; TISSUE=DEVELOPING ENDOSPERM;
 RX PubMed=11929961;
 RA Lid S.E., Grulis D., Jung R., Lorentzen J.A., Ananiev E.,
 RA Chamberlin M., Niu X., Meeley R., Nichols S., Olsen O.A.;
 RT "The defective kernel 1 (dek1) gene required for aleurone cell
 RT development in the endosperm of maize grains encodes a membrane
 RT protein of the calpain gene superfamily."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:5460-5465(2002).
 DR EMBL: AY061804; AAL38187.1;
 SO SEQUENCE 2159 AA; 238985 MW; 971C1A709BA5F5D4 CRC64;

Query Match 33.3%; Score 47; DB 10; Length 2159;
 Best Local Similarity 36.4%; Pred. No. 79;
 Matches 8; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 2 IGAUXXXXXXVLTGTPALISM 23
 Db 821 LGWALSTIAVITGTPALISM 842

RESULT 4
 ID 08ZHF4 PRELIMINARY; PRT: 197 AA.
 AC 08ZHF4;

DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein YP00945.
 GN YP00945.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Yersinia.
 OC NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Tiliakli R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Fellwell T., Hamlin N., Holroyd S., Jagsals K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague."
 RL Nature 413:523-527(2001).
 DR EMBL: AJ414145; CAC89788.1;
 DR InterPro: IPR002637; Ham1p-like.
 DR Pfam: PF01725; Ham1p-like; 1.
 DR PRODOM: PD004952; Ham1p-like; 1.
 DR TIGRFAMs: TIGR00042; Ham1p-like; 1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 197 AA; 20906 MW; 153FE882A5FDA639 CRC64;

Query Match 31.9%; Score 45; DB 16; Length 197;
 Best Local Similarity 52.9%; Pred. No. 14;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 4 AVLKXXXXVLTGTPAL 20
 Db 50 AILKARMAQTGTPAI 66

RESULT 5
 ID 0983X4 PRELIMINARY; PRT: 320 AA.

AC 0983X4:
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein mlr8126.
 GN MLR8126.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OC NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti."
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP003013; BAB53756.1;
 DR InterPro: IPR001296; Glycosyltransferase.
 DR Pfam: PF00534; Glycosyltransferase; 1.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 320 AA; 36354 MW; 68E97778C3815CA1 CRC64;

Query Match 31.9%; Score 45; DB 16; Length 320;
 Best Local Similarity 63.6%; Pred. No. 23;
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 13 LTTGLPALISM 23
 DB 103 LTTGMPRLIGW 113

RESULT 6
 Q31016 PRELIMINARY: PRT: 357 AA.

ID Q31016: 01-NOV-1996 (TREMblrel. 01, Created)
 AC Q31016: 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMblrel. 21, Last annotation update)
 DE MHC class I antigen SO-16 (Fragment).
 OS Saguinus oedipus (Cotton-top tamarin).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
 OX NCBI_TaxID=9490;

SEQUENCE FROM N.A.
 RP MEDLINE=91153864; PubMed=1840570;
 RA Watkins D.I., Garber T.L., Chen Z.W., Toukatey G., Hughes A.L.,
 RA Levin N.L.;
 RT "Unusually limited nucleotide sequence variation of the expressed
 RT major histocompatibility complex class I genes of a New World primate
 RT species (Saguinus oedipus).";
 RL Immunogenetics 33:79-89(1991).
 RN [2]

SEQUENCE FROM N.A.
 RA Watkins D.I., Garber T.L.;
 RL Submitted (Feb-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
 CC IMMUNE SYSTEM (BY SIMILARITY).
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN) (BY SIMILARITY).
 CC EMBL: U49331; AAA92140.1; -;
 DR HSSP: P30685; IAB8.
 DR InterPro: IPR003597; I9-cl.
 DR InterPro: IPR003006; I9_MHC.
 DR InterPro: IPR001039; MHC_I.
 DR Pfam: PF00047; I9; 1.
 DR Pfam: PF00129; MHC_I; 1.
 DR PRINTS: PR01638; MHCCLASSI.
 DR PRODOM: PD000050; MHC_I; 1.
 DR SMART: SM00407; IGcl; 1.
 DR PROSITE: PS00290; IG_MHC; 1.
 DR Glycoprotein; Transmembrane.
 FT NON_TER 1
 FT 357 1
 SO SEQUENCE 357 AA; 39969 MW; E7075EFF2BEA6A23 CRC64;

Query Match 31.9%; Score 45; DB 7; Length 357;
 Best Local Similarity 29.7%; Pred. No. 26;
 Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

OY 1 GIGAVLXXXXVLTGLPALISWIKXXXXXXRQ 37
 DB 302 GIVALLAILGAVTGAVVAVMRRKSSDKKGSYSQ 338

RESULT 7
 Q31013 PRELIMINARY: PRT: 357 AA.

ID Q31013: 01-NOV-1996 (TREMblrel. 01, Created)
 AC Q31013: 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMblrel. 21, Last annotation update)
 DE MHC class I Safu-G*02 (Fragment).
 GN MHC CLASS I SAFU-G.
 OS Saguinus fuscicollis (Brown-headed tamarin).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
 OX NCBI_TaxID=9487;

SEQUENCE FROM N.A.
 RP MEDLINE=98070787; PubMed=9405648;
 RA Cadavid L.F., Shuffeleboham C., Ruiz F.J., Yeager M., Hughes A.L.,
 RA Watkins D.I.;
 RT "Evolutionary instability of the major histocompatibility complex
 RT class I loci in new world primates.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:14536-14541(1997).
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
 CC IMMUNE SYSTEM (BY SIMILARITY).
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN) (BY SIMILARITY).
 CC EMBL: U59634; AAB97477.1; -;
 DR HSSP: O95352; IHHK.
 DR InterPro: IPR003597; I9-cl.
 DR InterPro: IPR003006; I9_MHC.
 DR InterPro: IPR001039; MHC_I.
 DR Pfam: PF00047; I9; 1.
 DR Pfam: PF00129; MHC_I; 1.
 DR PRINTS: PR01638; MHCCLASSI.
 DR PRODOM: PD000050; MHC_I; 1.
 DR SMART: SM00407; IGcl; 1.
 DR PROSITE: PS00290; IG_MHC; 1.
 DR Glycoprotein; Transmembrane.
 FT NON_TER 1
 FT 40347 1
 SO SEQUENCE 357 AA; 40347 MW; CDBA0A14B127F25D CRC64;

Query Match 31.9%; Score 45; DB 7; Length 357;
 Best Local Similarity 29.7%; Pred. No. 26;
 Matches 11; Conservative 5; Mismatches 21; Indels 0; Gaps 0;

OY 1 GIGAVLXXXXVLTGLPALISWIKXXXXXXRQ 37
 DB 302 GIVALLAILGAVTGAVVAVMRRKSSDKKGSYSQ 338

RESULT 8

Q31012 PRELIMINARY: PRT: 358 AA.

ID Q31012: 01-NOV-1996 (TREMblrel. 01, Created)
 AC Q31012: 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMblrel. 21, Last annotation update)
 DE MHC class I Safu-G*01 (Fragment).
 GN MHC SAFU-G.
 OS Saguinus fuscicollis (Brown-headed tamarin).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
 OX NCBI_TaxID=9487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98070787; PubMed=9405648;
 RA Cadavid L.F., Shuffeleboham C., Ruiz F.J., Yeager M., Hughes A.L.,
 RA Watkins D.I.;
 RT "Evolutionary instability of the major histocompatibility complex
 RT class I loci in new world primates.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:14536-14541(1997).
 CC -1- FUNCTION: INVOLVED IN THE PRESENTATION OF FOREIGN ANTIGENS TO THE
 CC IMMUNE SYSTEM (BY SIMILARITY).
 CC -1- SUBUNIT: DIMER OF ALPHA CHAIN AND A BETA CHAIN (BETA-2-
 CC MICROGLOBULIN) (BY SIMILARITY).
 CC EMBL: U59633; AAB97476.1; -;
 DR HSSP: O95352; IHHK.
 DR InterPro: IPR003597; I9-cl.
 DR InterPro: IPR003006; I9_MHC.
 DR InterPro: IPR001039; MHC_I.
 DR Pfam: PF00047; I9; 1.
 DR Pfam: PF00129; MHC_I; 1.
 DR PRINTS: PR01638; MHCCLASSI.
 DR PRODOM: PD000050; MHC_I; 1.
 DR SMART: SM00407; IGcl; 1.
 DR PROSITE: PS00290; IG_MHC; 1.
 DR Glycoprotein; Transmembrane.

Query Match 31.2% Score 44; DB 17; Length 457;
 Best Local Similarity 31.8%; Pred. No. 50;
 Matches 7; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

OY 3 GAVLKXXXXVLTGTPALISWI 24
 DB 281 GALKGSYIMLEGLPSFAGMM 302

RESULT 12

ID 018105 PRELIMINARY; PRT: 714 AA.

AC 018105;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
 DE T21B4.10 protein.
 GN T21B4.10.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Pelodertinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smye R.;
 RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RT none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RT Science 282:2012-2018(1998).
 DR EMBL; 281124; CAB03377.1;
 DR InterPro; IPR003003; 7TM_chemo2.
 DR InterPro; IPR000168; 7TM_nematode.
 DR Pfam; PF01604; 7tm_5; 2.
 SQ SEQUENCE 714 AA; 81848 MW; C3A1E2C1DF4PDBCE CRC64;

Query Match 31.2% Score 44; DB 5; Length 714;
 Best Local Similarity 40.9%; Pred. No. 79;
 Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 2 IGAVLKXXXXVLTGTPALISW 23
 DB 238 IGVLQYSPILLVGLTVAWS 259

RESULT 13

ID 09RY43 PRELIMINARY; PRT: 749 AA.

AC 09RY43;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Hypothetical protein DR0108.
 GN DR0108.
 OS Deinococcus radiodurans.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RL;
 RA MEDLINE=20036896; PubMed=10567266;
 RA White O., Eissen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Uterback T., Zalewski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;

RT "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1."
 RT Science 286:1571-1577(1999).
 DR EMBL; AF001873; AAF09696.1; -.
 DR TIGR; DR0108; -.

DR InterPro; IPR000731; HMGCR/patch_5TM.
 DR InterPro; IPR004869; MMP.
 DR Pfam; PF03176; MMP1; 2.
 DR PROSITE; PSS0156; SSD; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 749 AA; 79149 MW; 27947FD0F9D14A5 CRC64;

Query Match 31.2% Score 44; DB 16; Length 749;
 Best Local Similarity 50.0%; Pred. No. 83;
 Matches 12; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

OY 1 GIGAVLKXXXXVLT--TGLPALIS 22
 DB 315 GIGVLAVVLTVALSLTALPALIT 338

RESULT 14

ID 09A7G7 PRELIMINARY; PRT: 762 AA.

AC 09A7G7;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 DE Beta-D-glucosidase.
 GN CCI1756.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eissen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty K.,
 RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;

Complete genome sequence of Caulobacter crescentus.
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AEO05849; AAK23732.1; -.
 DR TIGR; CCI1756; -.
 DR InterPro; IPR002772; GH_3C.
 DR InterPro; IPR001764; GH_3N.
 DR InterPro; IPR002016; Peroxidase.
 DR Pfam; PF00933; Glyco_hydro_3; 1.
 DR Pfam; PF01915; Glyco_hydro_3_C; 1.
 DR PRINTS; PR00133; GHYDRILASE3.
 DR PROSITE; PSS00436; PEROXIDASE_2; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 762 AA; 81011 MW; CA1CC7B0D7E02DB CRC64;

Query Match 31.2% Score 44; DB 16; Length 762;
 Best Local Similarity 39.1%; Pred. No. 85;
 Matches 9; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 2 IGAVLKXXXXVLTGTPALISWI 24
 DB 516 VSAANKNAVVLLETGSPVLPMP 538

RESULT 15

ID 098K31 PRELIMINARY; PRT: 2078 AA.

AC 098K31;
 DT 01-OCT-2001 (TREMBLrel. 18, Created)

```

DT 01-OCT-2001 (TrEMBLrel. 18, last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, last annotation update)
DE Hypothetical protein m11661.
CN M11661.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takouchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL: AP002998; BAB48983.1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 2078 AA: 207618 MW: 22574885.9848 CRC64;

Query Match 31.28; Score 44; DB 16; Length 2078;
Best Local Similarity 47.6%; Pred. No. 2.4e+02;
Matches 10; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 GCGAVLKXXXVLTGTPALV 21
DB 1627 GAGLGKGGCTVTTAGTPALV 1647

```

Search completed: May 9, 2003, 16:17:36
 Job time : 28.0784 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:07:45 ; Search time 38.1176 Seconds
(without alignments)
94.386 Million cell updates/sec

Title: US-09-851-422a-8

Perfect score: 128
Sequence: 1 GFIAFLCTKVLDFGIDKLIQLIEDKXX 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126	98.4	25	23	AAO21730
2	126	98.4	25	23	AAO21739
3	126	98.4	25	23	AAE18192
4	126	98.4	25	23	AAE18197
5	126	98.4	26	23	AAO21737
6	126	98.4	26	23	AAO21738
7	126	98.4	26	23	AAE18200
8	126	98.4	27	23	AAO21736
9	126	98.4	27	23	AAE18199
10	115.5	90.2	27	23	AAE18201

11	111.5	87.1	24	23	AAO21741
12	111.5	87.1	24	23	AAE18203
13	111.5	87.1	26	23	AAE18204
14	111.5	87.1	26	23	AAE18205
15	54	42.2	412	21	AAO6275
16	54	42.2	412	21	AAO52720
17	54	42.2	421	23	ABE92983
18	54	42.2	442	21	AAO62774
19	54	42.2	442	21	AAO52719
20	54	42.2	459	21	AAO62718
21	54	42.2	459	21	AAO52718
22	53	41.4	709	21	AAV44482
23	50	39.1	233	22	ABE69320
24	49	38.3	205	21	AAE25173
25	49	38.3	273	23	ABE38171
26	48.5	37.9	529	18	AAW20994
27	48	37.5	223	21	AAV74695
28	48	37.5	638	23	ABE93645
29	47	36.7	222	22	AAV72762
30	47	36.7	372	21	AAE10937
31	47	36.7	372	21	AAE10938
32	47	36.7	372	21	AAE10939
33	47	36.7	372	21	AAE10940
34	47	36.7	372	21	AAE10941
35	47	36.7	372	21	AAE10942
36	47	36.7	372	21	AAE10943
37	47	36.7	372	21	AAE10944
38	47	36.7	717	23	ABE91481
39	47	36.7	806	23	ABE93653
40	47	36.7	984	23	ABE93326
41	46	35.9	131	19	AAW98363
42	46	35.9	161	21	AAE25410
43	46	35.9	481	23	AAO93169
44	46	35.9	828	23	ABE91804
45	46	35.9	913	22	ABE67486

ALIGNMENTS

RESULT 1
AAO21730
ID AAO21730 standard; Peptide: 25 AA.
XX
AC AAO21730:
DT 13-SEP-2002 (first entry)
XX
DE Amoebapore helix 3 cytolytic peptide.
XX
KW Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;
KW Ovary; prostate; breast; skin; lung; pancreas; Amoebapore helix 3.
XX
OS Entamoeba histolytica.
XX
PN US2002045736-A1.
PD 18-APR-2002.
XX
PF 27-AUG-2001; 2001US-0938623.
XX
PR 09-MAY-2001; 2001US-0851422.
XX
PA (YUXX/) YU X.
PA (WAGN/) WAGNER T E.
XX
PI YU X, Wagner TE;
XX
DR WPI: 2002-507251/54.
XX
PT A new procytotoxin useful in the treatment of cancer of e.g. prostate,
PT ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator
via a peptide bond cleavable by a specific protease

XX Disclosure: Page 3; 21pp; English.
PS The invention relates to a procytotoxin comprising a cytotoxic peptide
XX bound to an inactivator via a peptide bond, where the peptide bond is
CC susceptible to cleavage by a targeting specific protease. The
CC procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, breast, skin, lung or pancreas. This sequence represents the
CC Amebapore helix 3 cytolytic peptide relating to the invention.
XX
SQ Sequence 25 AA:
Query Match 98.4%; Score 126; DB 23; Length 25;
Best Local Similarity 100.0%; Pred. No. 6.1e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GFATLCTKVLDPGIDKLIQLIEDK 25
DB 1 GFATLCTKVLDPGIDKLIQLIEDK 25
RESULT 2
AAO21739 standard: peptide; 25 AA.
AC AAO21739;
DT 13-SEP-2002 (first entry)
XX Procytotoxin cytolytic peptide #4.
XX Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;
KM ovary; prostate; breast; skin; lung; pancreas; cytolytic.
XX
OS Undisclosed.
XX
PM Key Location/Qualifiers
FT Modified-site 9
FT /note- "This Lys residue is modified by R, where R is
FT independently selected from the group consisting of the
FT epsilon-amino group of the adjacent lysine residue,
FT [epsilon]-gamma-Glu, [epsilon]-gamma-Glu-[alpha]-gamma-
FT (Glu)1-3, [epsilon]-alpha-(Trp)1-3, [epsilon]-alpha-
FT (Lys)1-3, [epsilon]-alpha-(Arg)1-3, wherein [epsilon]-
FT gamma represents a peptide bond between the epsilon
FT amino group of lysine and the gamma carboxyl group of the
FT adjacent glutamate, [alpha]-gamma represents a peptide
FT bond between the alpha amino group of the first glutamate
FT and the gamma carboxyl group of the second glutamate,
FT [epsilon]-alpha represents a peptide bond between the
FT epsilon-amino acid of lysine and the alpha carboxyl group
FT of the indicated amino acid"
FT
FT Modified-site 17
FT /note- "This Lys residue is modified by R, where R is
FT independently selected from the group consisting of the
FT epsilon-amino group of the adjacent lysine residue,
FT [epsilon]-gamma-Glu, [epsilon]-gamma-Glu-[alpha]-gamma-
FT (Glu)1-3, [epsilon]-alpha-(Trp)1-3, [epsilon]-alpha-
FT (Lys)1-3, [epsilon]-alpha-(Arg)1-3, wherein [epsilon]-
FT gamma represents a peptide bond between the epsilon
FT amino group of lysine and the gamma carboxyl group of the
FT adjacent glutamate, [alpha]-gamma represents a peptide
FT bond between the alpha amino group of the first glutamate
FT and the gamma carboxyl group of the second glutamate,
FT [epsilon]-alpha represents a peptide bond between the
FT epsilon-amino acid of lysine and the alpha carboxyl group
FT of the indicated amino acid"
FT
FT Modified-site 25
FT /note- "This Lys residue is modified by R, where R is
FT independently selected from the group consisting of the
FT epsilon-amino group of the adjacent lysine residue,
FT [epsilon]-gamma-Glu, [epsilon]-gamma-Glu-[alpha]-gamma-
FT (Glu)1-3, [epsilon]-alpha-(Trp)1-3, [epsilon]-alpha-
FT (Lys)1-3, [epsilon]-alpha-(Arg)1-3, wherein [epsilon]-
FT gamma represents a peptide bond between the epsilon
FT amino group of lysine and the gamma carboxyl group of the
FT adjacent glutamate, [alpha]-gamma represents a peptide
FT bond between the alpha amino group of the first glutamate
FT and the gamma carboxyl group of the second glutamate,
FT [epsilon]-alpha represents a peptide bond between the
FT epsilon-amino acid of lysine and the alpha carboxyl group
FT of the indicated amino acid"

FT [epsilon]-gamma-Glu, [epsilon]-gamma-Glu-[alpha]-gamma-
FT (Glu)1-3, [epsilon]-alpha-(Trp)1-3, [epsilon]-alpha-
FT (Lys)1-3, [epsilon]-alpha-(Arg)1-3, wherein [epsilon]-
FT gamma represents a peptide bond between the epsilon
FT amino group of lysine and the gamma carboxyl group of the
FT adjacent glutamate, [alpha]-gamma represents a peptide
FT bond between the alpha amino group of the first glutamate
FT and the gamma carboxyl group of the second glutamate,
FT [epsilon]-alpha represents a peptide bond between the
FT epsilon-amino acid of lysine and the alpha carboxyl group
FT of the indicated amino acid"
FT
FT US2002045736-A1.
FT
FT 18-APR-2002.
FT
FT 27-AUG-2001; 2001US-0938623.
FT
FT 09-MAY-2001; 2001US-0851422.
FT
FT (YUXX/) YU X.
FT PA (MAGN/) MAGNER T E.
FT PI Yu X, Wagner TE;
FT DR WPI; 2002-507251/54.
FT
FT A new procytotoxin useful in the treatment of cancer of e.g. prostate,
FT ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator
FT via a peptide bond cleavable by a specific protease -
FT
FT Disclosure: Page 7; 21pp; English.
XX
XX The invention relates to a procytotoxin comprising a cytotoxic peptide
CC bound to an inactivator via a peptide bond, where the peptide bond is
CC susceptible to cleavage by a targeting specific protease. The
CC procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, breast, skin, lung or pancreas. This sequence represents a
CC procytotoxin cytolytic peptide sequence relating to the invention.
XX
SQ Sequence 25 AA:
Query Match 98.4%; Score 126; DB 23; Length 25;
Best Local Similarity 100.0%; Pred. No. 6.1e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GFATLCTKVLDPGIDKLIQLIEDK 25
DB 1 GFATLCTKVLDPGIDKLIQLIEDK 25
RESULT 3
AAE18192
ID AAE18192 standard: peptide; 25 AA.
XX
XX AAE18192;
DT 07-MAY-2002 (first entry)
XX Entamoeba histolytica procytotoxin, amebapore helix 3.
XX Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
KW prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX Entamoeba histolytica.
OS WO200185777-A2.
XX
XX 15-NOV-2001.
XX PD
XX 09-MAY-2001; 2001WO-US40690.
XX

PR	09-MAY-2000; 2000US-203063P.
PR	16-JUN-2000; 2000US-212042P.
XX	
XX	(GREE-) GREENVILLE HOSPITAL SYSTEM.
PA	
XX	
PI	Yu X, Wagner TE;
XX	
DR	WPI; 2002-154423/20.
XX	
PT	A procytotoxin comprising a cytotoxic peptide with at least one lysine
PT	residue which is converted into a cytotoxin in a target cell is used
PT	for treating cancer, particularly of the prostate, skin, ovary or lung
PT	-
XX	
PS	Disclosure; Page 7; 33pp; English.
XX	
CC	The present invention relates to a procytotoxin, comprising a cytotoxic
CC	peptide with at least one lysine residue bound by a peptide bond to at
CC	least one amino acid via the epsilon-amino group of the lysine residue.
CC	The procytotoxin is used to treat cancer, particularly of the prostate,
CC	ovary, lung or skin. Cytolytic peptide also known as channel-forming
CC	peptide, typically disrupt cell membrane, causing cell lysis and death
CC	(apoptosis) upon contact. The present sequence is Entamoeba histolytica
CC	procytotoxin, amoebapore helix 3.
XX	
SO	Sequence 25 AA:
	Query Match . 98.4%; Score 126; DB 23; Length 25;
	Best Local Similarity 100.0%; Pred. No. 6.1e-13;
	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GFIAFLCTKRVDFGIDKLIQLIEDK 25
DB	1 GFIAFLCTKRVDFGIDKLIQLIEDK 25
RESULT 4	
AAE18197	
ID	AAE18197 standard; peptide: 25 AA.
XX	
AC	AAE18197;
XX	
DT	07-MAY-2002 (first entry)
XX	
DE	Procytotoxin #1 for treating cancer.
KW	Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
KW	prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX	
OS	Unidentified.
XX	
FT	Key
FT	Modified-site
FT	9 Location/Qualifiers
FT	/note- "lys(R), where R is the unmodified epsilon-amino
FT	group of the adjacent lys, or is [epsilon-gamma]-Glu,
FT	[epsilon-gamma]-Glu-[alpha-gamma]-(Glu)1-3,
FT	[epsilon-alpha]-(Phe)1-3, [epsilon-alpha]-(Tyr)1-3,
FT	[epsilon-alpha]-(Trp)1-3, [epsilon-alpha]-(Ile)1-3 or
FT	[epsilon-alpha]-(Arg)1-3; where [epsilon-gamma]
FT	represents a peptide bond between the epsilon-amino
FT	group of lysine and the gamma carboxyl group of the
FT	adjacent glutamate, [alpha-gamma] represents a peptide
FT	bond between the alpha-amino group of the first glutamate
FT	and the gamma carboxyl group of the second glutamate,
FT	[epsilon-alpha] represents a peptide bond between the
FT	epsilon-amino acid of lysine and the alpha carboxyl
FT	group of the indicated amino acid, and 1-3 indicates that
FT	1, 2 or 3 of the designated amino acid can be linked to
FT	the first via conventional peptide bonds"
FT	17
FT	/note- "lys(R), where R is the unmodified epsilon-amino
FT	group of the adjacent lys, or is [epsilon-gamma]-Glu,
FT	[epsilon-gamma]-Glu-[alpha-gamma]-(Glu)1-3,

FT FT [lepsiion-alpha]-(Phe)1-3, [lepsiion-alpha]-(Tyr)1-3, [lepsiion-alpha]-(Arg)1-3, [lepsiion-alpha]-(Lys)1-3 or [lepsiion-alpha]-(Arg)1-3; where [lepsiion-gamma] represents a peptide bond between the epsilon amino group of lysine and the gamma carboxyl group of the adjacent glutamate, [alpha-gamma] represents a peptide bond between the alpha amino group of the first glutamate and the gamma carboxyl group of the second glutamate, [lepsiion alpha] represents a peptide bond between the epsilon amino acid of lysine and the alpha carboxyl group of the indicated amino acid, and 1-3 indicates that 1, 2 or 3 of the designated amino acid can be linked to the first via conventional peptide bonds"

25
 /note- "Lys(R), where R is the unmodified epsilon-amino group of the adjacent Lys, or is [lepsiion-gamma]-Glu, [lepsiion-gamma]-Glu-[alpha-gamma]-(Glu)1-3, [lepsiion-alpha]-(Phe)1-3, [lepsiion-gamma]-(Tyr)1-3, [lepsiion-alpha]-(Trp)1-3, [lepsiion-alpha]-(Lys)1-3 or [lepsiion-alpha]-(Arg)1-3; where [lepsiion-gamma] represents a peptide bond between the epsilon amino group of lysine and the gamma carboxyl group of the adjacent glutamate, [alpha-gamma] represents a peptide bond between the alpha amino group of the first glutamate and the gamma carboxyl group of the second glutamate, [lepsiion alpha] represents a peptide bond between the epsilon amino acid of lysine and the alpha carboxyl group of the indicated amino acid, and 1-3 indicates that 1, 2 or 3 of the designated amino acid can be linked to the first via conventional peptide bonds"

WO200185777-A2.
 15-NOV-2001..
 09-MAY-2001; 2001WO-US40690.
 09-MAY-2000; 2000US-203063P.
 16-JUN-2000; 2000US-212042P.
 (GREE-) GREENVILLE HOSPITAL SYSTEM.
 Yu X, Wagner TE;
 WPI; 2002-154423/20.

A procytotoxin comprising a cytotoxic peptide with at least one lysine residue which is converted into a cytotoxin in a target cell is used for treating cancer, particularly of the prostate, skin, ovary or lung

Claim 5; Page 23; 33pp; English.

The present invention relates to a procytotoxin, comprising a cytotoxic peptide with at least one lysine residue bound by a peptide bond to at least one amino acid via the epsilon-amino group of the lysine residue. The procytotoxin is used to treat cancer, particularly of the prostate, ovary, lung or skin. Cytolytic peptide also known as channel-forming peptide, typically disrupt cell membrane, causing cell lysis and death (apoptosis) upon contact. The present sequence is procytotoxin.

Sequence 25 AA:
 Query Match 98.4%; Score 126; DB 23; Length 25;
 Best local Similarity 100.0%; Pred. No. 6.1e-13;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 GFIAFLCTKVLDFGIDKLIQLIEDK 25
 1 GFIAFLCTKVLDFGIDKLIQLIEDK 25

RESULT 5

ID	AAO21737	standard; Peptide: 26 AA.
XX	AAO21737:	
AC	13-SEP-2002	(first entry)
DT		
XX		
DE	Procytotoxin cytolytic peptide #2.	
XX	Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;	
KW	ovary; prostate; breast; skin; lung; pancreas; cytolitic.	
OS	Unidentified.	
XX		
FH	Key	Location/Qualifiers
FT	Modified-site	25
FT	/note= "The epsilon amino group of Lys is peptide bonded	
FT	to the alpha carboxyl group of Phe"	
FT	Modified-site	26
FT	/note= "The alpha carboxyl group of Phe is peptide	
PN	bonded to the epsilon amino group of Lys"	
XX	US2002045736-A1.	
PD	18-APR-2002.	
XX		
PE	27-AUG-2001; 2001US-0938623.	
XX		
PR	09-MAY-2001; 2001US-0851422.	
XX		
PA	(YUXX/) YU X.	
PA	(MAGN/) MAGNER T E.	
XX		
PI	Yu X, Wagner TE;	
DR	WPI; 2002-507251/54.	
XX		
PT	A new procytotoxin useful in the treatment of cancer of e.g. prostate,	
PT	ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator	
PT	via a peptide bond cleavable by a specific protease	
XX		
PS	Disclosure; Page 7; 21pp; English.	
XX		
CC	The invention relates to a procytotoxin comprising a cytotoxic peptide	
CC	bound to an inactivator via a peptide bond, where the peptide bond is	
CC	susceptible to cleavage by a targeting specific protease. The	
CC	procytotoxin is used to treat cancer, particularly of the prostate,	
CC	ovary, breast, skin, lung or pancreas. This sequence represents a	
CC	procytotoxin cytolitic peptide sequence relating to the invention.	
XX		
SQ	Sequence	26 AA:
	Query Match	98.4%; Score 126; DB 23; Length 26;
	Best Local Similarity	100.0%; Pred. No. 6.4e-13;
	Matches 25; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GFATLCTKVLDPGIDKLQLIEDK 25	
DB	1 GFATLCTKVLDPGIDKLQLIEDK 25	
RESULT 6		
AAO21738		
ID	AAO21738 standard; Peptide: 26 AA.	
XX		
AC	AAO21738;	
XX		
DT	13-SEP-2002	(first entry)
XX		
DE	Procytotoxin cytolitic peptide #3.	
XX	Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;	
KW	ovary; prostate; breast; skin; lung; pancreas; cytolitic.	

XX	Unidentified.	Location/Qualifiers
OS		
XX		
FT	Key	17
FT	Modified-site	/note= "The epsilon amino group of Lys is peptide bonded
FT		to the alpha carboxyl group of Phe"
FT		25
FT	Modified-site	/note= "The epsilon amino group of Lys is peptide bonded
FT		to the alpha carboxyl group of Phe"
FT		26
FT	Modified-site	/note= "The alpha carboxyl group of Phe is peptide bonded
FT		to the epsilon amino group of Lys"
XX		
XX		
PN	US2002045736-A1.	
PD	18-APR-2002.	
XX		
PF	27-AUG-2001; 2001US-0938623.	
XX		
PR	09-MAY-2001; 2001US-0851422.	
XX		
PA	(YUXX/) YU X.	
XX	(WAGN/) WAGNER T E.	
PI	Yu X, Wagner TE;	
XX	WPI: 2002-507251/54.	
DR		
PT	A new procytotoxin useful in the treatment of cancer of e.g. prostate,	
PT	ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator	
PT	via a peptide bond cleavable by a specific protease	
PS	Disclosure: Page 7; 21pp: English.	
XX		
CC	The invention relates to a procytotoxin comprising a cytotoxic peptide	
CC	bound to an inactivator via a peptide bond, where the peptide bond is	
CC	susceptible to cleavage by a targeting specific protease. The	
CC	procytotoxin is used to treat cancer, particularly of the prostate,	
CC	ovary, breast, skin, lung or pancreas. This sequence represents a	
CC	procytotoxin cytolytic peptide sequence relating to the invention.	
XX		
SO	Sequence 26 AA;	
	Query Match 98.4%; Score 126; DB 23; Length 26;	
	Best Local Similarity 100.0%; Pred. NO. 6.4e-13;	
	Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1 GFATATCTKVLDFGIDKLIQLIEDK 25	
	1 GFIATICTKVLDFGIDKLIQLIEDK 25	
Db	1 GFATATCTKVLDFGIDKLIQLIEDK 25	
	RESULT 7	
	AAE18200	
ID	AAE18200 standard; peptide: 26 AA.	
XX		
AC	AAE18200;	
XX		
DT	07-MAY-2002 (first entry)	
XX		
DE	Proctoxin #1 for treating cancer.	
XX		
KW	Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;	
KM	prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.	
XX		
OS	Unidentified.	
XX		
FH	Key	Location/Qualifiers
FT	Modified-site	25
FT		/note= "Lys-[epsilon-alpha]; where [epsilon-alpha]
FT		represents a peptide bond between the epsilon amino
FT		group of lysine and the alpha carboxyl group of the

FT adjacent phenylalanine"
XX
XX WO200185777-A2.
XX
XX 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-US40690.
XX
XX 09-MAY-2000; 2000US-203063P.
PR 16-JUN-2000; 2000US-212042P.
XX
XX (GRE-) GREENVILLE HOSPITAL SYSTEM.
XX
XX Yu X, Wagner TE;
XX
XX WPI: 2002-154423/20.
XX
XX
XX A procytotoxin comprising a cytotoxic peptide with at least one lysine
PT residue which is converted into a cytotoxin in a target cell is used
PT for treating cancer, particularly of the prostate, skin, ovary or lung
PT
XX
XX
XX Disclosure; Page 12; 33pp; English.
XX
XX The present invention relates to a procytotoxin, comprising a cytotoxic
CC peptide with at least one lysine residue bound by a peptide bond to at
CC least one amino acid via the epsilon-amino group of the lysine residue.
CC The procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
CC peptide, typically disrupt cell membrane, causing cell lysis and death
CC (apoptosis) upon contact. The present sequence is procytotoxin.
XX
XX
SQ Sequence 26 AA:
Query Match 98.4%; Score 126; DB 23; Length 26;
Best Local Similarity 100.0%; Pred. No. 6.4e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GFATLCTKRVLDGDKLIQLIEDK 25
DB 1 GFATLCTKRVLDGDKLIQLIEDK 25
RESULT 8
AAO21736
ID AAO21736 standard; Peptide: 27 AA.
XX
XX AAO21736;
AC
XX 13-SEP-2002 (first entry)
DT
XX
XX Procytotoxin cytolytic peptide #1.
DE
XX
XX Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;
KW ovary; prostate; breast; skin; lung; pancreas; cytolytic.
XX
XX Unidentified.
OS
XX
XX
FH Key
FT Modified-site 25 Location/Qualifiers
FT Modified-site 25 /note= "The epsilon amino group of Lys is peptide bonded
FT Modified-site 26 to the gamma carboxyl group of Glu"
FT Modified-site 26 /note= "The gamma carboxyl group of this Glu is peptide
FT Modified-site 27 bonded to the epsilon amino group of Lys. The alpha amino
FT Modified-site 27 group of this Glu is peptide bonded to the gamma carboxyl
FT Modified-site 27 group of another Glu"
FT /note= "The gamma carboxyl group of this Glu is peptide
FT bonded to the alpha amino group of another Glu"
XX
XX US2002045736-A1.
XX

PD 18-APR-2002.
XX
XX 27-AUG-2001; 2001US-0938623.
XX
XX 09-MAY-2001; 2001US-0851422.
XX
XX (YUXX/) YU X.
PA (WAGN/) WAGNER T E.
XX
XX Yu X, Wagner TE;
XX
XX WPI: 2002-507251/54.
XX
XX
XX A new procytotoxin useful in the treatment of cancer of e.g. prostate,
PT ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator
PT via a peptide bond cleavable by a specific protease
PT
XX
XX Disclosure; Page 7; 21pp; English.
XX
XX The invention relates to a procytotoxin comprising a cytotoxic peptide
CC bound to an inactivator via a peptide bond, where the peptide bond is
CC susceptible to cleavage by a targeting specific protease. The
CC procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, breast, skin, lung or pancreas. This sequence represents a
CC procytotoxin cytolytic peptide sequence relating to the invention.
XX
XX
SQ Sequence 27 AA:
Query Match 98.4%; Score 126; DB 23; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.7e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 GFATLCTKRVLDGDKLIQLIEDK 25
DB 1 GFATLCTKRVLDGDKLIQLIEDK 25
RESULT 9
AAE18199
ID AAE18199 standard; peptide: 27 AA.
XX
XX AAE18199;
AC
XX 07-MAY-2002 (first entry)
DT
XX
XX Procytotoxin #3 for treating cancer.
DE
XX
XX Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
KW prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX
XX Unidentified.
OS
XX
XX
FH Key
FT Modified-site 25 Location/Qualifiers
FT Modified-site 25 /note= "Lys-[epsilon-gamma]; where [epsilon-gamma]
FT Modified-site 26 represents a peptide bond between the epsilon amino
FT Modified-site 26 group of lysine and the gamma carboxyl group of the
FT Modified-site 26 adjacent glutamate"
FT Modified-site 26 /note= "Glu-[alpha-gamma]; where [alpha-gamma]
FT Modified-site 26 represents a peptide bond between the alpha
FT Modified-site 26 amino group of the first glutamate and the gamma carboxyl
FT Modified-site 26 group of the second glutamate"
XX
XX WO200185777-A2.
XX
XX 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-US40690.
XX
XX 09-MAY-2000; 2000US-203063P.
PR 16-JUN-2000; 2000US-212042P.
XX

PA (GREG-) GREENVILLE HOSPITAL SYSTEM.
XX
PI Yu X, Wagner TE;
XX
DR WPI: 2002-154423/20.
XX
XX A procytotoxin comprising a cytotoxic peptide with at least one lysine
PT residue which is converted into a cytotoxin in a target cell is used
PT for treating cancer, particularly of the prostate, skin, ovary or lung
PT
PS Disclosure: Page 11: 33pp; English.
XX
XX The present invention relates to a procytotoxin, comprising a cytotoxic
CC peptide with at least one lysine residue bound by a peptide bond to at
CC least one amino acid via the epsilon-amino group of the lysine residue.
CC The procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
CC peptide, typically disrupt cell membrane, causing cell lysis and death
CC (apoptosis) upon contact. The present sequence is procytotoxin.
CC
XX
SO Sequence 27 AA:

Query Match 96.4%; Score 126; DB 23; Length 27;
Best Local Similarity 100.0%; Pred. No. 6,7e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GFIAFLCTKRVDFGIDKLIQLIEDK 25
Db 1 GFIAFLCTKRVDFGIDKLIQLIEDK 25

RESULT 10
AAE18201
ID AAE18201 standard; peptide: 27 AA.
XX
AC AAE18201;
XX
DT 07-MAY-2002 (first entry)
XX
DE Proctolin #2 for treating cancer.
XX
XX procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
KW prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
FT Modified-site 17..18
FT /note= "Lys([epsilon]-alpha)-Phe); where [epsilon]-alpha)
FT represents a peptide bond between the epsilon amino
FT group of lysine and the alpha carboxyl group of
FT phenylalanine, and a standard peptide linkage between
FT lysine and phenylalanine"
FT Modified-site 26
FT /note= "Lys-[epsilon]-alpha); where [epsilon]-alpha)
FT represents a peptide bond between the epsilon amino
FT group of lysine and the alpha carboxyl group of the
FT adjacent phenylalanine"
XX
XX WO200185777-A2.
XX
XX 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-USA0690.
XX
XX 09-MAY-2000; 2000US-203063P.
XX
XX 16-JUN-2000; 2000US-212042P.
XX
XX (GREG-) GREENVILLE HOSPITAL SYSTEM.
XX
XX
XX Yu X, Wagner TE;
XX

DR WPI: 2002-154423/20.
XX
XX A procytotoxin comprising a cytotoxic peptide with at least one lysine
PT residue which is converted into a cytotoxin in a target cell is used
PT for treating cancer, particularly of the prostate, skin, ovary or lung
PT
PS Disclosure: Page 12: 33pp; English.
XX
XX The present invention relates to a procytotoxin, comprising a cytotoxic
CC peptide with at least one lysine residue bound by a peptide bond to at
CC least one amino acid via the epsilon-amino group of the lysine residue.
CC The procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
CC peptide, typically disrupt cell membrane, causing cell lysis and death
CC (apoptosis) upon contact. The present sequence is prototoxin.
CC
XX
SO Sequence 27 AA:

Query Match 90.2%; Score 115.5; DB 23; Length 27;
Best Local Similarity 96.2%; Pred. No. 3.1e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 GFIAFLCTKRVDFGIDKLIQLIEDK 25
Db 1 GFIAFLCTKRVDFGIDKLIQLIEDK 26

RESULT 11
AAO21741
ID AAO21741 standard; peptide: 24 AA.
XX
AC AAO21741;
XX
DT 13-SEP-2002 (first entry)
XX
DE Procytotoxin cytolytic peptide #6.
XX
XX Cytotoxic; cytostatic; procytotoxin; inactivator; protease; cancer;
KW ovary; prostate; breast; skin; lung; pancreas; cytolitic.
XX
OS Unidentified.
XX
XX US2002045736-A1.
XX
XX 18-APR-2002.
XX
XX 27-AUG-2001; 2001US-0938623.
XX
XX 09-MAY-2001; 2001US-0851422.
XX
XX (YUXX/) YU X.
XX
XX (WAGN/) WAGNER T E.
XX
XX Yu X, Wagner TE;
XX
XX WPI: 2002-507251/54.
XX
XX A new procytotoxin useful in the treatment of cancer of e.g. prostate,
PT ovary, breast, or skin, has a cytotoxic peptide bound to an inactivator,
PT via a peptide bond cleavable by a specific protease
XX
XX Example 1: Page 10; 21pp; English.
XX
XX The invention relates to a procytotoxin comprising a cytotoxic peptide
CC bound to an inactivator via a peptide bond, where the peptide bond is
CC susceptible to cleavage by a targeting specific protease. The
CC procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, breast, skin, lung or pancreas. This sequence represents a
CC procytotoxin cytolitic peptide sequence relating to the invention.
XX
XX Sequence 24 AA:
XX

Query Match 87.1%; Score 111.5; DB 23; Length 24;
Best Local Similarity 96.0%; Pred. No. 1.2e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 GFATLCTKVLDFGIDKLIQLIEDK 25
|||||
DB 1 GFATLCTKVLDFGIDK-IQLIEDK 24

RESULT 12
AAE18203
ID AAE18203 standard; peptide: 24 AA.
XX
AC AAE18203;
XX
DT 07-MAY-2002 (first entry)
XX
DE Cytolytic peptide for treating cancer.
XX
KM Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
XX prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX
OS Unidentified.
XX
PN WO200185777-A2.
XX
PD 15-NOV-2001.
XX
PE 09-MAY-2001; 2001WO-US40690.
XX
PR 09-MAY-2000; 2000US-203063P.
XX
PT 16-JUN-2000; 2000US-212042P.
XX
PA (GREE-) GREENVILLE HOSPITAL SYSTEM.
XX
PI Yu X, Wagner TE;
XX
DR WPI: 2002-154423/20.
XX
PT A procytotoxin comprising a cytotoxic peptide with at least one lysine
PT residue which is converted into a cytotoxin in a target cell is used
PT for treating cancer, particularly of the prostate, skin, ovary or lung
PT
PS Example 1; Page 20; 33pp; English.
XX
CC The present invention relates to a procytotoxin, comprising a cytotoxic
CC peptide with at least one lysine residue bound by a peptide bond to at
CC least one amino acid via the epsilon-amino group of the lysine residue.
CC The procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
CC peptide, typically disrupt cell membrane, causing cell lysis and death
CC (apoptosis) upon contact. The present sequence is cytolytic peptide.
XX
SQ Sequence 24 AA;
XX

Query Match 87.1%; Score 111.5; DB 23; Length 24;
Best Local Similarity 96.0%; Pred. No. 1.2e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 GFATLCTKVLDFGIDKLIQLIEDK 25
|||||
DB 1 GFATLCTKVLDFGIDK-IQLIEDK 24

RESULT 13
AAE18204
ID AAE18204 standard; peptide: 26 AA.
XX
AC AAE18204;
XX
DT 07-MAY-2002 (first entry)
XX

DE Procytolytic peptide for treating cancer.
XX
KM Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
XX prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FH Modified-site 23
FT /note="Addition of two gamma glutamate linked side-chain
FT glutamate acid residue to the epsilon amino group of the
FT C-terminal lysine"

WO200185777-A2.
XX
PN 15-NOV-2001.
XX
PD 09-MAY-2001; 2001WO-US40690.
XX
PE 09-MAY-2000; 2000US-203063P.
XX
PR 16-JUN-2000; 2000US-212042P.
XX
PA (GREE-) GREENVILLE HOSPITAL SYSTEM.
XX
PI Yu X, Wagner TE;
XX
DR WPI: 2002-154423/20.
XX
PT A procytotoxin comprising a cytotoxic peptide with at least one lysine
PT residue which is converted into a cytotoxin in a target cell is used
PT for treating cancer, particularly of the prostate, skin, ovary or lung
PT
PS Example 1; Page 20; 33pp; English.
XX
CC The present invention relates to a procytotoxin, comprising a cytotoxic
CC peptide with at least one lysine residue bound by a peptide bond to at
CC least one amino acid via the epsilon-amino group of the lysine residue.
CC The procytotoxin is used to treat cancer, particularly of the prostate,
CC ovary, lung or skin. Cytolytic peptide also known as channel-forming
CC peptide, typically disrupt cell membrane, causing cell lysis and death
CC (apoptosis) upon contact. The present sequence is procytolytic peptide.
XX
SQ Sequence 26 AA;
XX

Query Match 87.1%; Score 111.5; DB 23; Length 26;
Best Local Similarity 96.0%; Pred. No. 1.3e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 GFATLCTKVLDFGIDKLIQLIEDK 25
|||||
DB 1 GFATLCTKVLDFGIDK-IQLIEDK 24

RESULT 14
AAE18205
ID AAE18205 standard; peptide: 26 AA.
XX
AC AAE18205;
XX
DT 07-MAY-2002 (first entry)
XX
DE Procytotoxin #5 for treating cancer.
XX
KM Procytotoxin; cytotoxic peptide; epsilon-amino group; cancer; cytostatic;
XX prostate; ovary; lung; skin; channel-forming peptide; apoptosis; therapy.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FH Modified-site 24.26
FT /note="Lys([epsilon-gamma]-Glu-[alpha-gamma]-Glu);
FT where [epsilon-gamma] represents a peptide bond between

FT the epsilon amino group of lysine and the gamma carboxyl
FT group of adjacent glutamate, [alpha-gamma] represents a
FT peptide bond between alpha amino group of glutamate and
FT gamma carboxyl group of adjacent glutamate"

XX WO200185777-A2.
XX
XX 15-NOV-2001.
XX
XX 09-MAY-2001; 2001WO-US40690.
XX
XX 09-MAY-2000; 2000US-203063P.
XX 16-JUN-2000; 2000US-212042P.
XX
XX (GREE-) GREENVILLE HOSPITAL SYSTEM.
XX
XX Yu X, Wagner TE;
XX
XX WPI; 2002-154423/20.
XX
XX A procytotoxin comprising a cytotoxic peptide with at least one lysine
XX residue which is converted into a cytotoxin in a target cell is used
XX for treating cancer, particularly of the prostate, skin, ovary or lung
XX
XX Claim 8; Page 24; 33pp; English.
XX
XX The present invention relates to a procytotoxin, comprising a cytotoxic
XX peptide with at least one lysine residue bound by a peptide bond to at
XX least one amino acid via the epsilon-amino group of the lysine residue.
XX The procytotoxin is used to treat cancer, particularly of the prostate,
XX ovary, lung or skin. Cytolytic peptide also known as channel-forming
XX peptide, typically disrupt cell membrane, causing cell lysis and death
XX (apoptosis) upon contact. The present sequence is procytotoxin.
XX
XX Sequence 26 AA:
XX
XX Query Match 87.1%; Score 111.5; DB 23; Length 26;
XX Best Local Similarity 96.0%; Pred. No. 1.3e-10;
XX Matches 24; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 GFIATLCTKVLDFGIDKLIQLIEDK 25
DB 1 GFIATLCTKVLDFGIDKLIQLIEDK 24

RESULT 15
AAG06275 standard; Protein: 412 AA.
XX
XX AAG06275;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 2989.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137122.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139493.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using SW model

Run on: May 9, 2003, 16:13:10 : Search time 9.79412 Seconds
(without alignments)
81.112 Million cell updates/sec

Title: US-09-851-422A-8

Sequence: 1 GFIAVLKRVLDGIDKLIQIEDKXX 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	39.1	499	4 US-08-426-509A-19	Sequence 19, Appl
2	50	39.1	499	5 PCT-US95-05008-19	Sequence 19, Appl
3	49	38.3	273	4 US-09-134-001C-3016	Sequence 3016, Ap
4	47	36.7	222	2 US-09-384-162-7	Sequence 7, Appl
5	46	35.9	54	2 US-08-456-647B-41	Sequence 41, Appl
6	46	35.9	54	2 US-08-237-401A-41	Sequence 41, Appl
7	46	35.9	279	2 US-08-701-191A-23	Sequence 23, Appl
8	46	35.9	984	2 US-08-673-789-9	Sequence 9, Appl
9	46	35.9	984	2 US-08-449-645A-19	Sequence 19, Appl
10	46	35.9	984	2 US-08-702-367A-19	Sequence 19, Appl
11	46	35.9	984	5 PCT-US95-04681-19	Sequence 19, Appl
12	44	34.4	259	2 US-09-134-001C-3640	Sequence 3640, Ap
13	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
14	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
15	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
16	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
17	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
18	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
19	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
20	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
21	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
22	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
23	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
24	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
25	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
26	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl
27	44	34.4	259	2 US-07-857-224B-54	Sequence 54, Appl

28	43	33.6	1082	1 US-08-357-598-10	Sequence 10, Appl
29	43	33.6	1082	2 US-09-003-289-5	Sequence 5, Appl
30	43	33.6	1082	2 US-09-003-289-10	Sequence 10, Appl
31	43	33.6	1082	5 PCT-US95-16435-5	Sequence 5, Appl
32	43	33.6	1082	5 PCT-US95-16435-10	Sequence 10, Appl
33	43	33.6	1124	4 US-09-191-786-1	Sequence 1, Appl
34	43	33.6	1455	2 US-08-726-012B-2	Sequence 2, Appl
35	42	32.8	675	4 US-08-426-509A-4	Sequence 4, Appl
36	42	32.8	675	5 PCT-US95-05008-4	Sequence 4, Appl
37	42	32.8	766	5 PCT-US94-00198-6	Sequence 6, Appl
38	42	32.8	1098	4 US-08-946-994-17	Sequence 17, Appl
39	42	32.8	1099	4 US-08-665-574C-16	Sequence 16, Appl
40	42	32.8	1100	1 US-08-357-598-11	Sequence 11, Appl
41	42	32.8	1100	2 US-09-003-289-11	Sequence 11, Appl
42	42	32.8	1100	5 PCT-US95-16435-11	Sequence 11, Appl
43	42	32.8	1299	5 PCT-US95-08354A-2	Sequence 2, Appl
44	41.5	32.4	274	2 US-08-701-191A-33	Sequence 33, Appl
45	41	32.0	76	2 US-08-469-537A-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1:
US-08-426-509A-19
Sequence 19, Application US/08426509A
Patent No. 6326469
GENERAL INFORMATION:
APPLICANT: Ullrich, Axel
APPLICANT: Gishizky, Mikhail
APPLICANT: Sures, Irman G.
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
TITLE OF INVENTION: TYROSINE KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESS: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,509A
FILING DATE: 21-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/232,545
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-0074-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 499 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: No. 6326469e
US-08-426-509A-19
Query Match 39.1%; Score 50; DB 4; Length 499;
Best Local Similarity 53.3%; Pred. No. 7;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,647B
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/237,401
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678-
REFERENCE/DOCKET NUMBER: 07251/007002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 54 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-647B-41

Query Match 35.9%; Score 46; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 2.2;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 LCTKVLDPGIDKLI 19
DB 8 LCKKVSDFGILTRL 21

RESULT 6
US-08-237-401A-41
Sequence 41, Application US/08237401A
Patent No. 5637448
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/237,401A
FILING DATE: 02-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Haile Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07251/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:

LENGTH: 54 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-237-401A-41

Query Match 35.9%; Score 46; DB 2; Length 54;
Best Local Similarity 57.1%; Pred. No. 2.2;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 LCTKVLDPGIDKLI 19
DB 8 LCKKVSDFGILTRL 21

RESULT 7
US-08-701-191A-23
Sequence 23, Application US/08701191A
Patent No. 5942428
GENERAL INFORMATION:
APPLICANT: Moosa Mohammadi, Joseph Schlesinger,
and Stevan R. Hubbard
TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
OF NON-INSULIN RECEPTOR TYROSINE KINASE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,191A
FILING DATE: August 21, 1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 227/088
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 279 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-701-191A-23

Query Match 35.9%; Score 46; DB 2; Length 279;
Best Local Similarity 57.1%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 LCTKVLDPGIDKLI 19
DB 151 LCKKVSDFGILTRL 164

RESULT 8
US-08-673-789-9
Sequence 9, Application US/08673789
Patent No. 5814479
GENERAL INFORMATION:
APPLICANT: ZHOU, REMPING; SCHULZ, NICHOLAS,
APPLICANT: T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
APPLICANT: GEORGE, F.
TITLE OF INVENTION: BSK RECEPTOR LIKE
TITLE OF INVENTION: TYROSINE KINASE AND LIGAND AND THEIR
TITLE OF INVENTION: USE IN DIAGNOSTIC AND THERAPEUTIC
METHODS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/673.789
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/177,812
FILING DATE: 04-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 984
TYPE: AMINO ACID
STRANDEDNESS: UNKNOWN
TOPOLOGY: UNKNOWN
US-08-673-789-9

Query Match 35.9%; Score 46; DB 2; Length 984;
Best Local Similarity 57.1%; Pred. No. 64;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 6 LCTKVLDFGIDKLI 19
DB 769 LCKKVSDFGLTRL 782

RESULT 9
US-08-449-645A-19
Sequence 19, Application US/08449645A
Patent No. 5981245
GENERAL INFORMATION:
APPLICANT: FOX, GARY M.
TITLE OF INVENTION: EPH-LIKE RECEPTOR PROTEIN TYROSINE
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN PATENT OPERATIONS/RBW
STREET: 1840 DEHAVILLAND DRIVE
CITY: THOUSAND OAKS
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 91320

COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,645A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WINTER, ROBERT B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 984 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-449-645A-19

Query Match 35.9%; Score 46; DB 2; Length 984;
Best Local Similarity 57.1%; Pred. No. 64;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 6 LCTKVLDFGIDKLI 19
DB 769 LCKKVSDFGLTRL 782

RESULT 10
US-08-702-367A-19
Sequence 19, Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
APPLICANT: FOX, GARY M.
TITLE OF INVENTION: EPH-LIKE RECEPTOR PROTEIN TYROSINE
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: AMGEN PATENT OPERATIONS/RBW
STREET: 1840 DEHAVILLAND DRIVE
CITY: THOUSAND OAKS
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,367A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: WINTER, ROBERT B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 984 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-702-367A-19

Query Match 35.9%; Score 46; DB 2; Length 984;
Best Local Similarity 57.1%; Pred. No. 64;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 6 LCTKVLDFGIDKLI 19
DB 769 LCKKVSDFGLTRL 782

Db 769 LCKKVSDFGLTRL 782

RESULT 11

PCT-US95-04681-19

; Sequence 19, Application PC/TUS9504681

; GENERAL INFORMATION:

; APPLICANT: Fox, Gary M.

; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine

; TITLE OF INVENTION: Kinases

; NUMBER OF SEQUENCES: 28

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Amgen Patent Operations/RBW

; STREET: 1840 Dehavilland Drive

; CITY: Thousand Oaks

; STATE: California

; COUNTRY: USA

; ZIP: 91320

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/04681

; FILING DATE:

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Winter, Robert B.

; REFERENCE/DOCKET NUMBER: A-287

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 984 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US95-04681-19

Query Match 35.9%; Score 46; DB 5; Length 984;

Best Local Similarity 57.1%; Pred. No. 64;

Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 LCTKVLDFGIDRLI 19

Db 769 LCKKVSDFGLTRL 782

RESULT 12

US-09-134-001C-3640

; Sequence 3640, Application US/09134001C

; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 3640

; LENGTH: 148

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

Query Match

Best Local Similarity 34.4%; Score 44; DB 4; Length 148;

Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 12 DFQIDKLIQIED 24

Db 48 DLGIDLVKRIKD 60

RESULT 13

US-07-857-224B-54

; Sequence 54, Application US/07857224B

; Patent No. 5958784

; GENERAL INFORMATION:

; APPLICANT: Benner, Steven A.

; TITLE OF INVENTION: Predicting Folded Structures of Proteins

; NUMBER OF SEQUENCES: 114

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Steven A. Benner

; STREET: Hadlaubstrasse 151

; CITY: Zurich

; STATE: none

; COUNTRY: Switzerland

; ZIP: (note: this is an international post code) CH-8092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage

; OPERATING SYSTEM: Apple Macintosh

; SOFTWARE: Microsoft Word

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/857,224B

; FILING DATE: 03/25/92

; CLASSIFICATION: 436

; PRIOR APPLICATION DATA: none

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (International) 41 1 632 2830

; TELEFAX: (International) 41 1 262 2437

; TELEX: none

; INFORMATION FOR SEQ ID NO: 54:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 259

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; DESCRIPTION: protein

; ORIGINAL SOURCE:

; ORGANISM: human

; FEATURE: Protein kinase; Table 8 Column 61

; PUBLICATION INFORMATION:

; AUTHORS:

; AUTHORS: Hanks, S. K.

; AUTHORS: Quinn, A. M.

; AUTHORS: Hunter, T.

; TITLE: The protein kinase family

; JOURNAL: Science

; VOLUME: 241

; PAGES: 42-52

; DATE: 1988

US-07-857-224B-54

Query Match 34.4%; Score 44; DB 2; Length 259;

Best Local Similarity 50.0%; Pred. No. 28;

Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 5 TLCTKVLDFGIDRLI 20

Db 133 TLCKIADFLARLIE 148

RESULT 14

US-07-857-224B-55

; Sequence 55, Application US/07857224B

; Patent No. 5958784

; GENERAL INFORMATION:

; APPLICANT: Benner, Steven A.

; TITLE OF INVENTION: Predicting Folded Structures of Proteins

NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven A. Benner
STREET: Hadlaubstrasse 151
CITY: Zurich
STATE: none
COUNTRY: Switzerland
ZIP: (note: this is an international post code) CH-8092
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/857,224B
FILING DATE: 03/25/92
CLASSIFICATION: 436
PRIOR APPLICATION DATA: none
TELECOMMUNICATION INFORMATION:
TELEPHONE: (international) 41 1 632 2830
TELEFAX: (international) 41 1 262 2437
TELEX: none
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 259
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: protein
ORIGINAL SOURCE:
ORGANISM: mouse
FEATURE: protein kinase; Table 8 Column 62
PUBLICATION INFORMATION:
AUTHORS:
AUTHORS: Hanks, S. K.
AUTHORS: Quinn, A. M.
AUTHORS: Hunter, T.
TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1988
US-07-857-224B-55

Query Match 34.4%; Score 44; DB 2; Length 259;
Best Local Similarity 50.0%; Pred. NO. 28;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 5 TLCTVVLDFGIDKLIQ 20
|||:||||:||||:
Db 133 TLSCKIADFGIARLIE 148

RESULT 15
US-09-457-040B-41
Sequence 41, Application US/09457040B
Patent No. 6387641
GENERAL INFORMATION:
APPLICANT: Vertex Pharmaceuticals Incorporated
APPLICANT: Bellon, Steve
TITLE OF INVENTION: Crystallized P38 Complexes
FILE REFERENCE: VPI/98-14
CURRENT APPLICATION NUMBER: US/09/457,040B
CURRENT FILING DATE: 1999-12-08
NUMBER OF SEQ ID NOS: 41
SOFTWARE: Patentin version 3.0
SEQ ID NO 41
LENGTH: 271
TYPE: PRT
ORGANISM: Human
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: x-lysine

US-09-457-040B-41
Query Match 34.4%; Score 44; DB 4; Length 271;
Best Local Similarity 50.0%; Pred. NO. 29;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 5 TLCTVVLDFGIDKLIQ 20
|||:||||:||||:
Db 145 TLSCKIADFGIARLIE 160

Search completed: May 9, 2003, 16:19:29
Job time : 11.7941 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:15:46 : Search time 11.3824 Seconds
(without alignments)
218.294 Million cell updates/sec

Title: US-09-851-422A-8
Refrect score: 128
Sequence: 1 GFIAITCTKVLDFGIDKLIQIEDKXX 27

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 349150 seqs, 92025710 residues

Total number of hits satisfying chosen parameters: 349150

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEM_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCIT_NEM_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEM_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEM_PUB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEM_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEM_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEM_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	39.1	499	9 US-09-977-260-19	Sequence 19, Appl
2	50	39.1	499	9 US-09-977-261-19	Sequence 19, Appl
3	50	39.1	499	10 US-09-977-269-19	Sequence 19, Appl
4	49	38.3	205	9 US-10-101-464A-492	Sequence 492, App
5	46	35.9	54	9 US-09-158-722-41	Sequence 41, Appl
6	46	35.9	131	10 US-09-881-752A-326	Sequence 326, App
7	46	35.9	161	9 US-10-101-464A-729	Sequence 729, App
8	45	35.2	717	10 US-09-801-368-376	Sequence 376, App
9	45	35.2	454	10 US-09-771-161A-95	Sequence 95, Appl
10	45	35.2	485	10 US-09-741-669-299	Sequence 299, App
11	45	35.2	505	10 US-09-771-161A-186	Sequence 186, App
12	45	35.2	827	9 US-10-101-464A-915	Sequence 915, App
13	44	34.4	122	9 US-10-101-464A-912	Sequence 912, App
14	44	34.4	346	10 US-09-870-962-3	Sequence 3, Appl
15	44	34.4	509	9 US-09-977-260-18	Sequence 18, Appl
16	44	34.4	509	9 US-09-977-261-18	Sequence 18, Appl
17	44	34.4	509	10 US-09-977-269-18	Sequence 18, Appl
18	44	34.4	1033	9 US-10-127-516-2	Sequence 2, Appl
19	44	34.4	1033	9 US-10-027-629-2	Sequence 2, Appl

20	43	33.6	169	9 US-10-101-464A-562	Sequence 562, App
21	43	33.6	309	9 US-10-101-464A-602	Sequence 602, App
22	43	33.6	347	10 US-09-815-242-10754	Sequence 10754, A
23	43	33.6	674	9 US-10-086-464-14	Sequence 14, Appl
24	43	33.6	968	9 US-10-101-464A-76	Sequence 76, Appl
25	43	33.6	984	9 US-10-101-464A-919	Sequence 919, App
26	43	33.6	1124	10 US-09-771-161A-199	Sequence 199, App
27	42.5	33.2	423	9 US-09-738-626-6367	Sequence 6367, Ap
28	42	33.8	172	10 US-09-810-264-26	Sequence 26, Appl
29	42	32.8	626	9 US-09-712-363-151	Sequence 151, App
30	42	32.8	675	9 US-10-186-399-3	Sequence 3, Appl
31	42	32.8	675	9 US-09-977-260-4	Sequence 4, Appl
32	42	32.8	675	9 US-09-977-261-4	Sequence 4, Appl
33	42	32.8	675	10 US-09-977-269-4	Sequence 4, Appl
34	42	32.8	724	9 US-09-881-579-4	Sequence 4, Appl
35	41	32.0	154	9 US-09-252-088-2	Sequence 2, Appl
36	41	32.0	188	9 US-09-738-628-5200	Sequence 5200, Ap
37	41	32.0	505	9 US-09-977-260-17	Sequence 17, Appl
38	41	32.0	505	9 US-09-977-261-17	Sequence 17, Appl
39	41	32.0	505	10 US-09-977-269-17	Sequence 17, Appl
40	40	32.0	928	9 US-10-108-605-261	Sequence 261, App
41	41	32.0	3672	10 US-09-927-668-12	Sequence 12, Appl
42	41	32.0	3801	10 US-09-927-668-10	Sequence 10, Appl
43	40.5	31.6	192	9 US-10-101-482-13	Sequence 13, Appl
44	40.5	31.6	331	10 US-09-033-525-2	Sequence 2, Appl
45	40	31.2	119	10 US-09-815-242-11367	Sequence 11367, A

ALIGNMENTS

RESULT 1
US-09-977-260-19
Sequence 19, Application US/09977260
Publicatation No. US20020192790A1
GENERAL INFORMATION:
APPLICANT: ULLRICH, AXEL
APPLICANT: GISHIZKY, MIKHAIL
APPLICANT: SURES, IRMINGARD
TITLE OF INVENTION: NOVEL MEKAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1260
CURRENT APPLICATION NUMBER: US/09/977,260
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 08/232,545
PRIOR FILING DATE: 1994-04-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 19
LENGTH: 499
TYPE: PRT
ORGANISM: Mus sp.
US-09-977-260-19
Query Match 39.1% Score 50; DB 9; Length 499;
Best Local Similarity 53.3% Pred. No. 8.8;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
DB 365 TLCTKVLDFGIDKLI 19
||| | :||| :|||
QY 5 TLCTKVLDFGIDKLI 19
DB 365 TLCTKVLDFGIDKLI 19
RESULT 2
US-09-977-261-19
Sequence 19, Application US/09977261
Publicatation No. US20030054527A1
GENERAL INFORMATION:
APPLICANT: ULLRICH, AXEL
APPLICANT: GISHIZKY, MIKHAIL
APPLICANT: SURES, IRMINGARD
TITLE OF INVENTION: NOVEL MEKAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1259
CURRENT APPLICATION NUMBER: US/09/977,261

; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-977-261-19

Query Match 39.1%; Score 50; DB 9; Length 499;
Best Local Similarity 53.3%; Pred. No. 8.8;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 5 TLCTKVLDFGIDKLI 19
||| | : ||| : ||| :
DB 365 TLCKKIADFGIARII 379

RESULT 3
US-09-977-269-19
; Sequence 19, Application US/09977269
; Patent No. US20020082037A1
; GENERAL INFORMATION:
; APPLICANT: ULBRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,269
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-977-269-19

Query Match 39.1%; Score 50; DB 10; Length 499;
Best Local Similarity 53.3%; Pred. No. 8.8;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 5 TLCTKVLDFGIDKLI 19
||| | : ||| : ||| :
DB 365 TLCKKIADFGIARII 379

RESULT 4
US-10-101-464A-492
; Sequence 492, Application US/10101464A
; Publication No. US20030046728A1
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and their use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 492
; LENGTH: 205
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-492

Query Match 38.3%; Score 49; DB 9; Length 205;
Best Local Similarity 56.2%; Pred. No. 4.6;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 ATLCCKVDFGIDKLI 19
||| | : ||| : ||| :
DB 86 ABFCPKVADFGIARIV 101

RESULT 5
US-09-158-722-41
; Sequence 41, Application US/09158722
; Publication No. US20030013848A1
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/158,722
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/456,647
; FILING DATE: 02-JUN-1995
; APPLICATION NUMBER: US 08/237,401
; FILING DATE: 02-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/884,486
; FILING DATE: 15-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Wettersell Ph.D., John R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: 07251/007002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 678-5070
; TELEFAX: (619) 678-5099
; INFORMATION FOR SEQ ID NO. 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 54 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-158-722-41

Query Match 35.9%; Score 46; DB 9; Length 54;
Best Local Similarity 57.1%; Pred. No. 2.9;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 6 LCTKVLDFGIDKLI 19
||| | : ||| : ||| :
DB 8 LCCKVSDFGIARIL 21

[illegible]

```

; Sequence 3, Application US/09870962
; Patent No. US20020081290A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzal, Yalda
; APPLICANT: Lu, Alina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/870,962
; CURRENT FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 09/420,915
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: US 09/173,581
; PRIOR FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 3
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 507669
US-09-870-962-3

```

```

Query Match          34.4%  Score 44; DB 10; Length 346;
Best Local Similarity 50.0%  Pred. No. 50;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 5 TLCTKVLDFGIDKLIQ 20
|| | : || : || :
Db 212 TLSCKIADFGIARLIE 227

```

```

RESULT 15
US-09-977-260-18
; Sequence 18, Application US/09977260
; Publication No. US20020192790A1
; GENERAL INFORMATION:
; APPLICANT: ULIRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,260
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-977-260-18

```

```

Query Match          34.4%  Score 44; DB 9; Length 509;
Best Local Similarity 50.0%  Pred. No. 77;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
QY 5 TLCTKVLDFGIDKLIQ 20
|| | : || : || :
Db 375 TLSCKIADFGIARLIE 390

```

Search completed: May 9, 2003, 16:20:18
 Job time : 12.3824 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:08:20 ; Search time 8.73529 Seconds

(without alignments)
128.200 Million cell updates/sec

Title: US-09-851-422a-8

Perfect score: 128
Sequence: 1 GFIAITCTKVLDFGIDKLIQLIEDKXX 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126	98.4	98	1	PPPA_ENTHI
2	125	97.7	97	1	PPPN_ENTHI
3	87	68.0	96	1	PPPB_ENTHI
4	76	59.4	101	1	PPPC_ENTHI
5	54	42.2	521	1	SPKA_SYNY3
6	53	41.4	835	1	OBP_VZVD
7	50	38.1	498	1	BLK_MOUSE
8	49	38.3	507	1	LCK_CHICK
9	48.5	37.9	559	1	PRIM_HELPJ
10	47.5	37.1	559	1	PRIM_HELPJ
11	47	36.7	290	1	VG33_HSV11
12	46	35.9	535	1	GCP_METUA
13	46	35.9	738	1	STIL_YEAST
14	46	35.9	976	1	EPAL_HUMAN
15	45.5	35.5	660	1	Y390_MYCPN
16	45	35.2	445	1	YXEO_BACSU
17	45	35.2	485	1	YGF0_ECOLI
18	45	35.2	504	1	BLK_HUMAN
19	45	35.2	657	1	GRAD_TREPA
20	44.5	34.8	372	1	PRS2_METTH
21	44.5	34.8	441	1	YV07_METUA
22	44	34.4	261	1	YGS1_YEAST
23	44	34.4	379	1	Y829_BRUME
24	44	34.4	508	1	LCK_HUMAN
25	44	34.4	508	1	LCK_MOUSE
26	44	34.4	1062	1	NAL2_MOUSE
27	43	33.6	66	1	SECE_RICPR
28	43	33.6	491	1	ZAPA_PROMI
29	43	33.6	1124	1	JAK3_HUMAN
30	43	33.6	1455	1	FACA_HUMAN
31	42.5	33.2	304	1	Y202_METUA
32	42.5	33.2	423	1	BIOA_CORGL
33	42	32.8	65	1	SECE_STACA

34	42	32.8	163	1	ATPX_OCHNE
35	42	32.8	395	1	MERC_ECOLI
36	42	32.8	503	1	HCK_RAT
37	42	32.8	504	1	HCK_MACFA
38	42	32.8	524	1	HCK_MOUSE
39	42	32.8	622	1	PKNB_MYCLE
40	42	32.8	626	1	PKNB_MYCTU
41	42	32.8	675	1	BMX_HUMAN
42	42	32.8	693	1	PKNL_MYXHA
43	42	32.8	766	1	GAP1_SCHPO
44	42	32.8	966	1	VIA_BMV
45	42	32.8	1100	1	JAK3_RAT

ALIGNMENTS

RESULT 1
ID PPPA_ENTHI STANDARD: PRT: 98 AA.
AC P34095; Q24839; (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pore-forming peptide. amebapore A precursor (EH-APP).
OC Entamoeba histolytica.
OS Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-46.
RC STRAIN-HM-1:IMSS;
RX MEDLINE=93010939; PubMed=1396552;
RA Leippe M., Tannich E., Nickel R., van der Goot G., Patus F.,
Horsmann R.D., Mueller-Eberhard H.J.;
RT Primary and secondary structure of the pore-forming peptide of
RT pathogenic Entamoeba histolytica.*;
RL EMBO J. 11:3501-3506(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-HM-1:IMSS;
RX MEDLINE=94098992; PubMed=8274224;
RA Bruchhaus I., Leippe M., Lioutas C., Tannich E.;
RT "Unusual gene organization in the protozoan parasite Entamoeba
RT histolytica.*";
RL DNA Cell Biol. 12:925-933(1993).
RN [3]
RP SEQUENCE OF 22-46.
RX MEDLINE=91352048; PubMed=1881907;
RA Leippe M., Ebel S., Schoenberger O.L., Horsmann R.D.,
Mueller-Eberhard H.J.;
RT "Pore-forming peptide of pathogenic Entamoeba histolytica.*";
RL Proc. Natl. Acad. Sci. U.S.A. 88:7659-7663(1991).
RN [4]
RP PARTIAL SEQUENCE.
RC STRAIN-HM-1:IMSS;
RX MEDLINE=95231296; PubMed=7715451;
RA Leippe M., Andrae J., Nickel R., Tannich E., Mueller-Eberhard H.J.;
RT "Amebapores, a family of membranolytic peptides from cytoplasmic
RT granules of Entamoeba histolytica: isolation, primary structure, and
RT pore formation in bacterial cytoplasmic membranes.*";
RL Mol. Microbiol. 14:895-904(1994).
CC -!- FUNCTION: FORMS PORES IN THE CYTOPLASMIC MEMBRANE OF HOST CELLS.
CC HAS ANTIBACTERIAL ACTIVITY AGAINST M. LUTEBUS, NO ACTIVITY AGAINST
CC E. COLI. IMPLICATED IN THE CYTOLYTIC ACTIVITY OF THE PARASITE.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES.
CC -!- SIMILARITY: CONTAINS 1 SAPOININ B-TYPE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@lsb-sib.ch.
CC -----
DR EMBL: M83945; AAA29111.1;
DR EMBL: X70851; CAA50203.1;
DR PIR: S25283; S25283.
DR PIR: A41279; A41279.
DR InterPro: IPR000004; SAPP.
DR SMART: SM00118; SAPP; 1.
KW Signal; Antibiotic.
FT SIGNAL 1 21
FT PEPTIDE 22 98 PORE-FORMING PEPTIDE AMEOBAPORE A.
FT DISULFID 26 98 BY SIMILARITY.
FT DISULFID 29 92 BY SIMILARITY.
FT DISULFID 56 67 BY SIMILARITY.
FT VARIANT 71 71 L->F.
SQ SEQUENCE 98 AA: 10504 MW: F9592FDC1PAC7FEA CRC64;

Query Match 98.4%; Score 126; DB 1; Length 98;
Best Local Similarity 100.0%; Pred. No. 3; Le-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GFIATLCTKVLDFGIDKLIQLIEDK 25
Db 61 GFIATLCTKVLDFGIDKLIQLIEDK 85
|||||

RESULT 2
ID PFPC_ENTHI STANDARD: PRT: 97 AA.
AC 007831;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Nonpathogenic pore-forming peptide precursor (APNP).
OS Entamoeba histolytica.
CC Eukaryota; Entamoebidae; Entamoeba.
CC NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-34.
RC STRAIN=SAW 142;
RX MEDLINE=93295428; PubMed=8515772;
RA Lelepe M., Bahr E., Tannich E., Horstmann R.D.;
RT "Comparison of pore-forming peptides from pathogenic and
RT nonpathogenic Entamoeba histolytica";
RL Mol. Biochem. Parasitol. 59:101-110(1993).
CC -1- FUNCTION: FORMS PORE IN THE MEMBRANE OF HOST TISSUES. IMPLICATED
CC IN THE CYTOLYTIC ACTIVITY OF THE PARASITE.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: CONTAINS 1 SAPOSIN B-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: L04984; AAA18632.1;
DR InterPro: IPR000004; SAPP.
DR SMART: SM00118; SAPP; 1.
KW Signal.
FT SIGNAL 1 20
FT CHAIN 21 97 NONPATHOGENIC PORE-FORMING PEPTIDE.
FT DISULFID 25 97 BY SIMILARITY.
FT DISULFID 28 91 BY SIMILARITY.
FT DISULFID 55 66 BY SIMILARITY.
SQ SEQUENCE 97 AA: 10432 MW: AE171C58B86E007 CRC64;

Query Match 97.7%; Score 125; DB 1; Length 97;
Best Local Similarity 96.0%; Pred. No. 4; Le-12;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GFIATLCTKVLDFGIDKLIQLIEDK 25
Db 60 GFIATLCTKVLDFGVDKLIQLIEDK 84
|||||

RESULT 3
ID PFPC_ENTHI STANDARD: PRT: 96 AA.
AC 024824;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pore-forming peptide amebopore B precursor (EH-APP).
OS Entamoeba histolytica.
CC Eukaryota; Entamoebidae; Entamoeba.
CC NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 20-63.
RC STRAIN=HM-1:IMSS;
RX MEDLINE=95231296; PubMed=7715451;
RA Lelepe M., Andrae J., Nickel R., Tannich E., Mueller-Eberhard H.J.;
RT "Amebopores, a family of membranolytic peptides from cytoplasmic
RT granules of Entamoeba histolytica: Isolation, primary structure, and
RT pore formation in bacterial cytoplasmic membranes.";
RL Mol. Microbiol. 14:895-904(1994).
CC -1- FUNCTION: FORMS PORES IN THE CYTOPLASMIC MEMBRANE OF HOST CELLS.
CC HAS ANTIBACTERIAL ACTIVITY AGAINST M. LUTEUS, NO ACTIVITY AGAINST
CC E. COLI. IMPLICATED IN THE CYTOLYTIC ACTIVITY OF THE PARASITE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES.
CC -1- SIMILARITY: CONTAINS 1 SAPOSIN B-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.lsb-sib.ch/announce/>
CC or send an email to license@lsb-sib.ch).
CC -----
DR EMBL: X76904; CAA54226.1;
DR InterPro: IPR000004; SAPP.
DR SMART: SM00118; SAPP; 1.
KW Signal; Antibiotic.
FT SIGNAL 1 19
FT PEPTIDE 20 96 PORE-FORMING PEPTIDE AMEOBAPORE B.
FT DISULFID 24 96 BY SIMILARITY.
FT DISULFID 27 90 BY SIMILARITY.
FT DISULFID 54 65 BY SIMILARITY.
SQ SEQUENCE 96 AA: 10437 MW: 99E6AFAD802B42CC CRC64;

Query Match 68.0%; Score 87; DB 1; Length 96;
Best Local Similarity 58.3%; Pred. No. 2; Le-06;
Matches 14; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Oy 1 GFIATLCTKVLDFGIDKLIQLIEDK 24
Db 59 GFIATLCTKVLDFGVDELKLIEN 82
|||||

RESULT 4
ID PFPC_ENTHI STANDARD: PRT: 101 AA.
AC 024825;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pore-forming peptide amebopore C precursor (EH-APP).
OS Entamoeba histolytica.
CC Eukaryota; Entamoebidae; Entamoeba.
CC NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-69.
RC STRAIN=HM-1:IMSS;


```

RX MEDLINE=95231296; PubMed=7715451.
RT "Amoebapores, a family of membranolytic peptides from cytoplasmic granules of Entamoeba histolytica: isolation, primary structure, and pore formation in bacterial cytoplasmic membranes."
RL Mol. Microbiol. 14:895-904(1994).
CC -1- FUNCTION: FORMS PORES IN THE CYTOPLASMIC MEMBRANE OF HOST CELLS. HAS ANTIBACTERIAL ACTIVITY AGAINST M. LUTEUS, NO ACTIVITY AGAINST E. COLI. IMPLICATED IN THE CYTOLYTIC ACTIVITY OF THE PARASITE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES.
CC -1- SIMILARITY: CONTAINS 1 SAPOSTEIN B-TYPE DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC
CC
DR EMBL: X76904; CA54225.1; -.
DR InterPro: IPR000004; SApB.
DR SMART: SM00118; SApB; 1.
DR Signal: Antibiotic.
FT SIGNAL 1 24
FT PEPTEID 25 101 PORE-FORMING PEPTIDE AMOEBAPORE C.
FT DISULFID 29 101 BY SIMILARITY.
FT DISULFID 32 95 BY SIMILARITY.
FT DISULFID 59 70 BY SIMILARITY.
SQ SEQUENCE 101 AA; 10855 MW; 28AC72135E734998 CRC64;
Query Match 59.4%; Score 76; DB 1; Length 101;
Best Local Similarity 59.1%; Pred. NO. 0.00012;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0.
OY 1 GFIAFLCTKRVLFDFGDKLIQLI 22
Db 64 GLVEFLCTKRVSYGDKLIERI 85
1 : ::::: ::::: :
64 : ::::: ::::: :

RESULT 5
SPKA_SYNY3 STANDARD: PRT: 521 AA.
AC QPAB3: P73208; P73209;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Probable serine/threonine-protein kinase A (EC 2.7.1.-).
GN SPKA OR SLI1574/SLI1575.
OS Synchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA Kamei A., Ikeuchi M.;
RT "A novel gene, SPKA, encodes active Ser/Thr protein kinase in the motile cyanobacterium Synchocystis sp. PCC 6803."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hitosawa M., Sugitara M., Sasamoto S., Kimura T., Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S., Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium Synchocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions."
RL DNA Res. 3:109-126(1996).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to a frameshift in position 103.
CC

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AB046597; BAB17033.1; -
CC DR EMBL; D90904; BAA17236.1; ALT.FRAME.
CC DR EMBL; D90904; BAA17235.1; ALT.FRAME.
CC DR InterPro; IPR000719; Euk_Pkinase.
CC DR InterPro; IPR002290; Ser_Thr_Pkinase.
CC DR InterPro; IPR001245; Tyr_Pkinase.
CC DR Pfam; PF00069; Pkinase.1.
CC DR PRINTS; PR00109; TYRKINASE.
CC DR ProDom; PD000001; Euk_Pkinase.1.
CC DR SMART; SM00220; S_TKc.1.
CC DR SMART; SM00219; TYTKc.1.
CC DR PROSITE; PS00107; PROTEIN_KINASE_ATP.1.
CC DR PROSITE; PS00108; PROTEIN_KINASE_ST.1.
CC DR PROSITE; PS50011; PROTEIN_KINASE_DOM.1.
CC KW Transferrase; Serine/threonine-protein kinase; ATP-binding;
CC KM Complete proteome.
CC FT DOMAIN 15 289 PROTEIN KINASE.
CC FT NP_BIND 21 29 ATP (BY SIMILARITY).
CC FT BINDING 45 45 ATP (BY SIMILARITY).
CC FT ACT_SITE 148 148 BY SIMILARITY.
CC SO SEQUENCE 521 AA; 58874 MW; 5FEC16F0DD369C91 CRC64;
OY Query Match 42.2%; Score 54; DB 1; Length 521;
Db Best Local Similarity 62.5%; Pred. No. 1.3;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 9 KVLDFGDKLTIQLIED 24
Db 167 KILDFGIKLVQAMEE 182
I:|||||I|I|I|I|:
-----
RESULT 6
OBP_VZVD
AC OBP_VZVD STANDARD: PRT; 835 AA.
AC P09293;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Replication origin binding protein (OBP).
DE 51.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
OX NCBI_TaxID=10338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1759-1816(1986).
CC -I- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION. BINDS THE ORIGIN
CC OF REPLICATION (ORI) (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL9,
CC EHV-1 53, AND VZV 51.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X04370; CAA27933.1; -

```

```

DR PIR: G27344; WZBES1.
DR InterPro: IPR003450; Herpes_orf_bp.
DR Pfam: PF02399; Herpes_orf_1.
DR DNA replication; DNA-binding; ATP-binding.
DR NP_BIND 67 74
DR BIND 67 74
DR SEQUENCE 835 AA; 94374 MW; A71F5877ACF386FB CRC64;

Query Match 41.4%; Score 53; DB 1; Length 835;
Best Local Similarity 43.5%; Pred. No. 3;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 1 GFIATLCTKVLDFGIDKLIQLE 23
DB 225 GFSGRCTILRDMGIDTLVRVTK 247

RESULT 7
BLK_MOUSE STANDARD: PRT; 498 AA.
ID BLK_MOUSE
AC P16277;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase BLK (EC 2.7.1.112) (B lymphocyte kinase) (p55-BLK).
DE BLK.
GN BLK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RE SEQUENCE FROM N.A.
RC TISSUE=B-cell;
RX MEDLINE=90117147; PubMed=2404338;
RA Dymecki S.M., Niederhuber J.E., Desiderio S.V.;
RT "Specific expression of a tyrosine kinase gene, blk, in B lymphoid
RT cells."
RL Science 247:332-336(1990).
RN [2]
RP STRUCTURE BY NMR OF SH2 DOMAIN.
RX MEDLINE=96224819; PubMed=8639560;
RA Metzler W.J., Letting B., Pryor K., Mueller L., Farmer B.T. II;
RT "The three-dimensional solution structure of the SH2 domain from
RT p55blk kinase."
RL Biochemistry 35:6201-6211(1996).
CC -1- FUNCTION: BLK MAY FUNCTION IN A SIGNAL TRANSDUCTION PATHWAY THAT
CC IS RESTRICTED TO B LYMPHOID CELLS.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
CC tyrosine phosphate.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M30903; AAA40453.1; -
DR PIR: A40092; A40092.
DR PDB: 1BLJ; 12-MAR-97.
DR PDB: 1BLK; 12-MAR-97.
DR MCD: MGI:88159; BLK.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF00018; SH3; 1.

```

```

DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRINTS: PR00452; SH3DOMAIN.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_kinase; 1.
DR ProDom: PD000066; SH3; 1.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR SMART: SM00326; SH3; 1.
DR SMART: SM00219; Tyrc; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_TYR; FALSE-NEG.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50002; SH3; 1.
DR Tyrosine-protein kinase; Phosphorylation; Transferase; ATP-binding;
KW Myristate; SH2 domain; SH3 domain; 3D-structure.
FT INIT_MET 0 0
FT LIPID 1 1
FT DOMAIN 51 111
FT DOMAIN 117 213
FT DOMAIN 234 487
FT NP_BIND 240 248
FT BINDING 262 262
FT ACT_SITE 353 353
FT MOD_RES 382 382
SQ SEQUENCE 498 AA; 56513 MW; BE49D7B079FDD577 CRC64;

Query Match 39.1%; Score 50; DB 1; Length 498;
Best Local Similarity 53.3%; Pred. No. 5.1;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 5 TLCTKVLDFGIDKLI 19
DB 364 TLCKKIDFGCLARIL 378

RESULT 8
LCK_CHICK STANDARD: PRT; 507 AA.
ID LCK_CHICK
AC P42683;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Proto-oncogene tyrosine-protein kinase LCK (EC 2.7.1.112) (Protein-
DE tyrosine kinase C-TKL).
GN Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RE SEQUENCE OF 1-88 FROM N.A.
RX MEDLINE=92186854; PubMed=1545804;
RA Chow L., Ratcliffe M., Veilleux A.;
RT "tkl is the avian homolog of the mammalian lck tyrosine protein
RT kinase gene."
RL Mol. Cell. Biol. 12:1226-1233(1992).
RN [2]
RP SEQUENCE OF 46-507 FROM N.A.
RX MEDLINE=88097370; PubMed=3321053;
RA Streibhardt K., Mullins J.I., Bruck C., Ruebsamen-Waigmann H.;
RT "Additional member of the protein-tyrosine kinase family: the src-
RT and lck-related protooncogene c-tkl."
RL Proc. Natl. Acad. Sci. U.S.A. 84:8778-8782(1987).
CC -1- FUNCTION: MAY PARTICIPATE IN ANTIGEN-INDUCED T-CELL ACTIVATION.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
CC tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: BOUND TO THE CYTOPLASMIC DOMAIN OF EITHER
CC CMA OR CD8.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC
CC SUBFAMILY.

```

```

CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M85043; AAA49003.1; -
CC EMBL: J03579; AAA49081.1; ALT_INIT.
CC HSSP: P06239; 31CK.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR000980; SH2.
CC InterPro: IPR001452; SH3.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00017; SH2; 1.
CC Pfam: PF00018; SH3; 1.
CC Pfam: PF00069; pkinase; 1.
CC PRINTS: PR00401; SH2DOMAIN.
CC PRINTS: PR00452; SH3DOMAIN.
CC PRINTS: PR00109; TYRKINASE.
CC ProDom: PD000001; Euk_pkinase; 1.
CC ProDom: PD000066; SH3; 1.
CC ProDom: PD000093; SH2; 1.
CC SMART: SM00252; SH2; 1.
CC SMART: SM00326; SH3; 1.
CC SMART: SM00219; TyKc; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE: PS50001; SH2; 1.
CC PROSITE: PS50002; SH3; 1.
CC Proto-oncogene: Tyrosine-protein kinase; Phosphorylation; Transferase;
CC ATP-binding; Myristate; SH2 domain; SH3 domain; Palmitate;
CC Lipoprotein.
CC INIT_MET 0 0 PROBABLE.
CC LIPID 1 1 MYRISTATE (BY SIMILARITY).
CC LIPID 2 2 PALMITATE (BY SIMILARITY).
CC LIPID 4 4 PALMITATE (BY SIMILARITY).
CC DOMAIN 59 119 SH2.
CC DOMAIN 125 222 SH3.
CC DOMAIN 243 496 PROTEIN KINASE.
CC NP_BIND 249 257 ATP (BY SIMILARITY).
CC BINDING 271 271 ATP (BY SIMILARITY).
CC ACT_SITE 362 362 BR SIMILARITY.
CC MOD_RES 392 392 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC MOD_RES 503 503 PHOSPHORYLATION (NEGATIVE REGULATION) (BY
CC SIMILARITY).
CC SEQUENCE 507 AA; 58008 MW; BC83CAFA691B6170 CRC64;

Query Match 38.3%; Score 49; DB 1; Length 507;
Best Local Similarity 53.3%; Pred. No. 7.3;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 6 LCTKVLDFGIDKLIQ 20
DB 374 LCKTADFGELARLIE 388

RESULT 9
PRIM_HELPY STANDARD; PRT; 559 AA.
AC Q92N49;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA primase (EC 2.7.7.-).
GN DNAG OR JHP0010.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;

```

```

OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moll D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tumano P.J., Carso A., Urie-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Voyis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 397:176-180(1999).
CC -1- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
CC RNA PRIMERS FOR THE OKAZAKI FRAGMENTS ON BOTH TEMPLATE STRANDS AT
CC REPLICATION FORKS DURING CHROMOSOMAL DNA SYNTHESIS.
CC -1- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE001441; AAD05592.1; -
CC HSSP: Q9X4D0; 1D0Q.
CC InterPro: IPR002936; DNAPrim_toprim.
CC InterPro: IPR002694; znf_CHC2.
CC Pfam: PF01751; Toprim; 1.
CC Pfam: PF01807; znf-CHC2; 1.
CC ProDom: PD002988; znf-CHC2; 1.
CC SMART: SM00493; TOPRIM; 1.
CC SMART: SM00400; znf-CHC2; 1.
CC Transferase; DNA replication; DNA-directed RNA polymerase; Primosome;
CC Zinc-finger; Zinc; Metal-binding; Complete proteome.
CC ZN_FING 37 61
CC ZN_FING 37 61 CHC2-TYPE (BY SIMILARITY).
CC SEQUENCE 559 AA; 63815 MW; 7B13BACAF94F87 CRC64;

Query Match 37.9%; Score 48.5; DB 1; Length 559;
Best Local Similarity 47.8%; Pred. No. 9.7;
Matches 11; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

OY 6 LCTKVLDFGIDKLIQ 21
DB 149 LCTNKIDYGHENKGNKDKLIQL 171

RESULT 10
PRIM_HELPY STANDARD; PRT; 559 AA.
AC P56064;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA primase (EC 2.7.7.-).
GN DNAG OR HP0012.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khairat H.G., Gilek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,

```

RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen *Helicobacter*
RT *pylori*."
RL Nature 368:539-547(1997).
CC -1- FUNCTION: DNA PRIMASE IS THE POLYMERASE THAT SYNTHESIZES SMALL
CC RNA PRIMERS FOR THE OKAZAKI FRAGMENTS ON BOTH TEMPLATE STRANDS AT
CC REPLICATION FORKS DURING CHROMOSOMAL DNA SYNTHESIS.
CC -1- COFACTOR: BINDS ONE ZINC ION PER MOLECULE (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE00523; AAD07082.1; -
CC HSSP: Q9X4D0; 1D0Q.
CC TIGR: HP0012; -
CC InterPro: IPR002936; DMAPrim_toprim.
CC InterPro: IPR002694; Znf_CHC2.
CC Pfam: PF01807; TOPrim_1.
CC Pfam: PF01807; zfc-CHC2; 1.
CC ProDom: PD002988; Znf-CHC2; 1.
CC SMART: SM00493; TOPrim; 1.
CC SMART: SM00400; Znf-CHC2; 1.
CC Transfaser: DNA replication; DNA-directed RNA polymerase; Primosome;
CC Zinc-finger; Zinc; Metal-binding; Complete proteome.
CC ZN_FING 37 61 CHC2-TYPE (BY SIMILARITY).
SQ SEQUENCE 559 AA; 63723 MW; E5EB7EF0B8B8170 CRC64;
Query Match 37.1%; Score 47.5; DB 1; Length 559;
Best Local Similarity 47.8%; Pred. No. 14;
Matches 11; Conservative 3; Mismatches 2; Indels 7; Gaps 1;
DB 149 LCTNRIDYGIENKGLINKDLIEL 171
111 :1:11 :11:1
6 LCTKVLDFGI-----DKLIQL 21
111 :1:11 :11:1
149 LCTNRIDYGIENKGLINKDLIEL 171
RESULT 11
VG33_HSV11 STANDARD: PRT; 290 AA.
AC 000118;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Hypothetical gene 33 protein.
GN 33.
OS Ictalurid herpesvirus 1 (Channel catfish virus) (CCV).
OC Viruses: dsDNA viruses, no RNA stage: Herpesviridae;
OC Ictalurid Herpes-like viruses.
OC NCBI_TaxId=10401;
RN NCB1
RP SEQUENCE FROM N.A.
RC STRAIN=Auburn 1;
RX MEDLINE=92087490; PubMed=1727613;
RA Davison A.J.;
RT "Channel catfish virus: a new type of herpesvirus."
RT Virology 186:9-14(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL: M75136; AAA68136.1; -
DR PIR: G36789; G36789.
DR Hypothetical protein.
KW SEQUENCE 290 AA; 34173 MW; 21C26260965A7E5E CRC64;
Query Match 36.7%; Score 47; DB 1; Length 290;
Best Local Similarity 75.0%; Pred. No. 8.3;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
DB 217 GFTTCTCKELD 228
111 :1:11 :11:1
1 GFTTCTCKELD 12
RESULT 12
GCP_METUA
ID GCP_METUA STANDARD: PRT; 535 AA.
AC 058350;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative O-sialoglycoprotein endopeptidase (EC 3.4.24.57)
DE (Glycoprotease).
DE MJI130.
GN Methanococcus jannaschii.
OS Archaea: Euryarchaeota: Methanococci: Methanococcales;
OC Methanocaldococcaceae: Methanocaldococcus.
OX NCBI_TaxId=2190;
RN NCB1
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.C., Fuhman J.L., Nguyen D.,
RA Usterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*
RT *jannaschii*."
RL Science 273:1058-1073(1996).
CC -1- FUNCTION: COULD BE A METALLOPROTEINASE.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of O-sialoglycoproteins: cleaves
CC 31-Arg-1-Asp-32 bond in glycoprotein A. Does not cleave
CC unglycosylated proteins, desialylated glycoproteins or
CC glycoproteins that are only N-glycosylated.
CC -1- COFACTOR: ZINC (PROBABLE).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M22.
CC -1- SIMILARITY: IN THE C-TERMINAL, SOME SIMILARITY WITH CONSERVED
CC REGION OF CATALYTIC DOMAIN OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U67555; AAB9132.1; -
CC MEROPS: M22.0PW; -
CC TIGR: MJI130; -
CC InterPro: IPR000719; Euk_pkkinase.
CC InterPro: IPR000905; Peptidase_M22.
CC InterPro: IPR001245; Tyr-phinase.
CC Pfam: PF00814; Peptidase_M22; 1.
CC PRINTS: PR00789; OSIALOPTASE.
CC ProDom: PD000001; Euk_pkkinase; 1.
CC ProDom: PD002367; Peptidase_M22; 1.
CC TIGRFAMS: TIGR00329; gcp; 1.
CC PROSITE: PS01016; GLYCOPROTEINASE; 1.

DR PROSITE: PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; UNKNOWN_1.
 KW Hypothetical protein; Hydrolase; Metalloprotease; Zinc;
 Complete proteome.
 FT DOMAIN 1 312 METALLOPEPTIDASE.
 FT DOMAIN 333 535 PROTEIN_KINASE_LINK.
 FT METAL 106 106 ZINC (POTENTIAL).
 FT METAL 110 110 ZINC (POTENTIAL).
 SO SEQUENCE 535 AA; 60570 MW; B18EEFACE8E5A99F CRC64;

Query Match 35.9%; Score 46; DB 1; Length 535;
 Best Local Similarity 50.0%; Pred. No. 22;
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 10 VLDGIDKLIQLEDK 25
 DB 465 IIDFGIKTSNDEK 480

RESULT 13
 ID STILL_YEAST STANDARD: PRT: 738 AA.
 AC P23561:
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Serine/threonine-protein kinase STE11 (EC 2.7.1.-).
 GN STE11 OR YLR362W OR I8039.10.
 OS Saccharomyces cerevisiae (baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 OX NCBI_TaxId=4932;
 RN [1]
 RP MEDLINE=9115076; PubMed=2276621;
 RX Rhodes N., Connell L., Errede B.;
 RT STE11 is a protein kinase required for cell-type-specific
 RT transcription and signal transduction in yeast.*;
 RL Genes Dev. 4:1862-1874(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Favallo A., Fulton L., Gattung S., Greco T., Kirsten J.,
 RA Kucaba T., Hallsworth K., Hawkins J., Hallier L., Jier M.,
 RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
 RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
 RA Rifken L., Riles L., Tatch A., Trevasakis E., Vignati D.,
 RA Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterston R.;
 RA Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP POSSIBLE FUNCTION.
 RX MEDLINE=92331935; PubMed=1628833;
 RA Cairns B.R., Ramey S.W., Kornberg K.D.;
 RT Order of action of components in the yeast pheromone response
 RT pathway revealed with a dominant allele of the STE11 kinase and the
 RT multiple phosphorylation of the STE7 kinase.*;
 RL Genes Dev. 6:1305-1318(1992).
 RN [4]
 RP FUNCTION: SERINE/THREONINE PROTEIN KINASE REQUIRED FOR CELL-TYPE-
 RP SPECIFIC TRANSCRIPTION AND SIGNAL TRANSDUCTION IN YEAST. IT
 RP IS THOUGHT THAT IT PHOSPHORYLATES THE STE7 PROTEIN KINASE WHICH
 RP ITSELF, PHOSPHORYLATES THE FUS3 AND OR KSS1 KINASES.
 CC -1- PATHWAY: RESPONSE TO PHEROMONE-INDUCED SIGNAL.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC -1- MAP KINASE KINASE SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL: X53431; CAA37522.1;
 CC EMBL: U19103; AAB67571.1;
 CC DR PIR: A36456; A36456.
 CC SCD: S0004354; STE11.
 CC InterPro: IPR000719; Euk_pkinase.
 CC InterPro: IPR001660; SAM.
 CC InterPro: IPR002290; Ser_thr_pkinase.
 CC Pfam: PF00069; pkinase; 1.
 CC Pfam: PF00536; SAM; 1.
 CC ProDom: PD000001; Euk_pkinase; 1.
 CC SMART: SM00454; SAM; 1.
 CC SMART: SM00220; S_TKC; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 CC PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC DR PROSITE: PS50105; SAM_DOMAIN; 1.
 CC Transferase: Serine/threonine-protein kinase; ATP-binding;
 KW Pheromone response.
 FT DOMAIN 41 105 SAM.
 FT NP_BIND 436 733 PROTEIN_KINASE.
 FT NP_BIND 442 450 ATP (BY SIMILARITY).
 FT BINDING 465 465 ATP (BY SIMILARITY).
 FT ACT_SITE 600 600 BY SIMILARITY.
 SQ SEQUENCE 738 AA; 83217 MW; A50D6980B346A41 CRC64;

Query Match 35.9%; Score 46; DB 1; Length 738;
 Best Local Similarity 42.1%; Pred. No. 31;
 Matches 8; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 7 CTRVLDGIDKLIQLEDK 25
 DB 613 CKNIDPFGIKTSNDEK 631

RESULT 14
 ID EPAL_HUMAN STANDARD: PRT: 976 AA.
 AC P21709; O15405;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Ephrin type-A receptor 1 precursor (EC 2.7.1.112) (Tyrosine-protein
 DE kinase receptor EPH).
 DE EPH1 OR EPH1 OR EPH1 OR EPH1 OR EPH1.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88070650; PubMed=2825356;
 RA Hirai H., Maru Y., Hagiwara K., Nishida J., Takaku F.;
 RT A novel putative tyrosine kinase receptor encoded by the eph gene.*;
 RL Science 238:1717-1720(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9299440; PubMed=10369740;
 RA Owsalimpur D., Kelley M.J.;
 RT Genomic structure of the EPH1 receptor tyrosine kinase gene.*;
 RL Mol. Cell. Probes 13:169-173(1999).
 RN [3]
 RP SEQUENCE OF 286-976 FROM N.A.
 RP TISSUE=Placenta;
 RA Tuzi N.L.;
 RT Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR MEMBERS OF THE EPHRIN-A FAMILY. BINDS WITH
 CC A LOW AFFINITY TO EPHRIN-A1.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: OVEREXPRESSED IN SEVERAL CARCINOMAS.

CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. EPHRIN RECEPTOR SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M18391; AAA36747.1; ALT_SEQ.
 DR EMBL: AF101171; AAD3440.1; JOINED.
 DR EMBL: AF101165; AAD3440.1; JOINED.
 DR EMBL: AF101166; AAD3440.1; JOINED.
 DR EMBL: AF101167; AAD3440.1; JOINED.
 DR EMBL: AF101168; AAD3440.1; JOINED.
 DR EMBL: AF101169; AAD3440.1; JOINED.
 DR EMBL: AF101170; AAD3440.1; JOINED.
 DR EMBL: Z27409; CA81796.1; JOINED.
 DR PIR: A34076; A34076.
 DR HSSP: P00523; 2PTK.
 DR Genew: HGN3385; EPH1.
 DR MIM: 179610; .
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR001090; Ephrin_receptor.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR InterPro: IPR001426; Kinase_receptor.
 DR Pfam: PF00041; fn3_2.
 DR Pfam: PF00069; pkinase_1.
 DR Pfam: PF00536; SAM_1.
 DR Pfam: PF01404; EPH_Lbd_1.
 DR PRINTS: PR00109; TYRKINASE.
 DR PRODOM: PD000001; Euk_pkinase_1.
 DR PRODOM: PD001495; Ephrin_receptor_1.
 DR SMART: SM00060; FN3_2.
 DR SMART: SM00454; SAM_1.
 DR SMART: SM00219; TYRK_1.
 DR PROSITE: PS01186; EGF_2; UNKNOWN_1.
 DR PROSITE: PS0107; PROTEIN_KINASE_ATP_1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM_1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR_1.
 DR PROSITE: PS00790; RECEPTOR_TYR_KIN_V_1.
 DR PROSITE: PS00791; RECEPTOR_TYR_KIN_V_2; 1.
 DR PROSITE: PS50105; SAM_DOMAIN_1.
 DR Transferrase: Tyrosine-protein kinase: ATP-binding; Phosphorylation;
 DR Receptor: Transmembrane; Glycoprotein; Signal; Repeat.
 FT SIGNAL 1 23
 FT CHAIN 24 976
 FT DOMAIN 24 547
 FT TRANSMEM 548 568
 FT DOMAIN 569 976
 FT DOMAIN 191 329
 FT DOMAIN 330 442
 FT DOMAIN 443 537
 FT DOMAIN 624 884
 FT DOMAIN 913 976
 FT SITE 974 976
 FT NP_BIND 630 638
 FT BINDING 656 656
 FT ACT_SITE 749 749
 FT MOD_RES 599 599
 FT MOD_RES 605 605
 FT MOD_RES 781 781
 FT MOD_RES 930 930
 FT CARBOHYD 338 338
 FT CARBOHYD 414 414
 FT CARBOHYD 478 478

FT CONFLICT 398 398 A -> G (IN REF. 1).
 FT CONFLICT 581 616
 FT
 FT SEQUENCE 976 AA: 106080 MW: 50512CCFL3F03808 CRC64;
 Query Match 35.9%; Score 46; DB 1; Length 976;
 Best Local Similarity 57.1%; Pred. No. 41;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 QY 6 LCTKVLDFGIDKLI 19
 DB 761 LCKKVSDFGLTRLL 774
 RESULT 15
 ID Y390_MYCPN STANDARD; PRT; 660 AA.
 AC P75207;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical ATP-binding protein MG390 homolog (D02_orf660).
 GN MPN571 OR MP271.
 OS Mycoplasma pneumoniae.
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelreich R., Hilbert H., Plagens H., Pickl E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae".
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -1- SIMILARITY: LIMITED, TO ABC TRANSPORTERS ATP-BINDING PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000027; AAB95919.1; .
 DR InterPro: IPR003439; ABC_transport.
 DR InterPro: IPR005074; Peptidase_C39.
 DR Pfam: PF00005; ABC_tran_1.
 DR Pfam: PF03412; Peptidase_C39_1.
 DR Hypothetical protein: ATP-binding; Complete proteome.
 FT NP_BIND 494 501
 FT SEQUENCE 660 AA: 76014 MW: B2A97A32F793A7 CRC64;
 QY 8 TKVLDFGI-----DKLIQLIEDK 25
 DB 422 TRIIOFGAANVLQFCRKDLVTLFEDK 448
 Query Match 35.5%; Score 45.5; DB 1; Length 660;
 Best Local Similarity 37.0%; Pred. No. 33;
 Matches 10; Conservative 4; Mismatches 4; Indels 9; Gaps 1;
 Search completed: May 9, 2003, 16:16:20
 Job time : 10.7353 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:10:40 ; Search time 18.5294 Seconds
(without alignments)
300.240 Million cell updates/sec

Title: US-09-851-422a-8
Perfect score: 128
Sequence: 1 GFATLCTKVLDFGIDKLIQIEDKXX 27

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protist:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriophage:*
17: sp_archaeop:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	68.8	96	5 09U8G5	09U8G5 entamoeba d
2	69	53.9	101	5 09U8G4	09U8G4 entamoeba d
3	63	49.2	87	5 015591	015591 entamoeba h
4	54	42.2	421	10 049409	049409 arabidopsis
5	54	41.4	459	10 093230	093230 arabidopsis
6	53	40.6	835	12 09J3M7	09J3M7 human herpe
7	52	40.6	396	17 026243	026243 methanobact
8	51	39.8	491	17 087V25	087V25 methanopyru
9	51	39.8	520	16 081NP6	081NP6 anabena sp
10	50	39.1	213	16 09A314	09A314 caulobacter
11	50	39.1	233	5 09V112	09V112 dirosophila
12	50	39.1	816	12 09E1W9	09E1W9 cercopithec
13	49	38.3	309	3 008743	008743 saccharomyc
14	49	38.3	629	10 0943D4	0943D4 oryza sativ
15	48.5	37.9	551	3 096X52	096X52 pholiota na
16	48	37.5	371	5 023326	023326 caenorhabdi

17	48	37.5	425	16 099RP5	099RP5 staphylococ
18	48	37.5	443	10 0943D1	0943D1 oryza sativ
19	48	37.5	481	16 P72727	P72727 synchocyst
20	48	37.5	638	10 09FF31	09FF31 arabidopsis
21	48	37.5	654	10 0943C7	0943C7 oryza sativ
22	48	37.5	1274	5 060985	060985 dictyostell
23	47.5	37.1	885	10 09LPC6	09LPC6 arabidopsis
24	47	36.7	131	16 092M33	092M33 helicobacte
25	47	36.7	183	16 092UM3	092UM3 rhizobium m
26	47	36.7	317	10 09M2J1	09M2J1 arabidopsis
27	47	36.7	317	5 09Y074	09Y074 leishmania
28	47	36.7	372	5 09Y081	09Y081 leishmania
29	47	36.7	372	5 09Y079	09Y079 leishmania
30	47	36.7	372	5 09U1G1	09U1G1 leishmania
31	47	36.7	372	5 09Y077	09Y077 leishmania
32	47	36.7	372	5 025260	025260 leishmania
33	47	36.7	372	5 09XTM3	09XTM3 leishmania
34	47	36.7	537	10 0943D6	0943D6 oryza sativ
35	47	36.7	818	10 0943D9	0943D9 oryza sativ
36	47	36.7	806	10 09FID5	09FID5 arabidopsis
37	47	36.7	984	10 09L2V1	09L2V1 arabidopsis
38	46.5	36.3	257	5 09U7C4	09U7C4 dictyostell
39	46.5	36.3	258	5 08T1N4	08T1N4 dictyostell
40	46.5	36.3	379	2 09ZEX1	09ZEX1 bordetella
41	46.5	36.3	859	16 09A6P4	09A6P4 caulobacter
42	46	35.9	131	16 025024	025024 helicobacte
43	46	35.9	215	16 09PHP9	09PHP9 campylobact
44	46	35.9	319	16 0988H0	0988H0 rhizobium l
45	46	35.9	386	16 08X2K5	08X2K5 ralsomonia s

ALIGNMENTS

RESULT 1	ID	Q9U8G5	PRELIMINARY:	PRT:	96 AA.
AC	Q9U8G5:				
DT	01-MAY-2000 (TREMBLrel. 13, Created)				
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE	Pore-forming protein isoform B precursor.				
GN	DP-B.				
OS	Entamoeba dispar.				
OC	Eukaryota; Entamoebidae; Entamoeba.				
NC	NCBI_TaxID=46681;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-SAW 142;				
RX	MEDLINE=99449604; PubMed=10518795;				
RA	Nickel R., Ott C., Dandekar T., Leippe M.,				
RT	"Pore-forming peptides of Entamoeba dispar similarity and divergence				
RT	to amoebapores in structure, expression and activity.";				
RL	Eur. J. Biochem. 265:1002-1007(1999).				
DR	EMBL: AF082528; AAF04195.1; -				
DR	InterPro: IPR000004; SAPB.				
DR	Prodom: PD001732; SAPB: 1.				
DR	SMART: SM00118; SAPB: 1.				
KW	SIGNAL.				
FT	CHAIN	1	19	POTENTIAL.	
FT	CHAIN	20	96	PORE-FORMING PROTEIN ISOFORM B.	
SQ	SEQUENCE	96 AA:	10480 MW;	9FA42AC9A9FAFB2 CRC64;	
Query Match			58.8%;	Score 88;	DB 5; Length 96;
Best Local Similarity			58.3%;	Pred. No. 5.7e-06;	
Matches 14; Conservative			7;	Mismatches 3;	Indels 0;
Gaps			0;	Gaps	0;
OY	1 GFATLCTKVLDFGIDKLIQIED 24				
DB	59 GFLGTLCNKILSFQVDELVKLIEN 82				
RESULT 2					

ID	Q9UG64	PRELIMINARY:	PT:	101 AA.
AC	Q9UG64;			
DT	01-MAY-2000 (TREMBLrel. 13. Created)			
DT	01-MAY-2000 (TREMBLrel. 13. Last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21. Last annotation update)			
DE	Pore-forming protein isoform C precursor.			
GN	DP-C.			
OS	Entamoeba dispar.			
OC	Eukaryota; Entamoebidae; Entamoeba.			
OX	NCBI_taxid=46681;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-SAW 142.			
RX	MEDLINE=99449604; PubMed=10518795;			
RA	Nickel R., Ott C., Dandekar T., Lelpe M.;			
RT	"Pore-forming peptides of entamoeba dispar similarity and divergence			
RT	to amoebozoes in structure, expression and activity.";			
RL	Eur. J. Biochem. 265:1002-1007(1999).			
DR	EMBL: AF082529; AAF04196.1; -.			
DR	InterPro: IPR000004; SApB.			
DR	ProDom: PD001732; SApB; 1.			
DR	SMART: SM00118; SApB; 1.			
KW	Signal.			
FT	SIGNAL. 1 24			POTENTIAL.
FT	CHAIN 25 101			PORE-FORMING PROTEIN ISOFORM C.
SQ	SEQUENCE 101 AA; 10877 MW; 7699A7EA9E0876EB CRC64;			

Query Match	53.98;	Score 69;	DB 5;	Length 101;
Best Local Similarity	54.50;	Pred. NO. 0.0053;		
Matches 12;	Conservative 6;	Mismatches 4;	Indels 0;	Caps 0;

```
QY      1 GFATLCTKVLDFGIDKLIQLI 22
          | | | | | | | | | |
Db      64 GVIEIIVCSKIVSYGIDKLIKI 85
```

RESULT 3	.
015591	
ID 015591	PRELIMINARY;
	PRT;
	87 AA

01-JUN-1998 (FRENBlrel. 05, created)
 DT 01-JUN-1998 (FRENBlrel. 05, last sequence update)
 DT 01-MAR-2002 (FRENBlrel. 20, last annotation update)
 DE Anopheles C. homologue (Fragment).
 OS Entamoeba histolytica.
 OC Eukaryota; Eucaryotidae; Entamoeba.
 NCBI_TaxID=5759;

RP	SEQUENCE FROM N.A.
RN	11

RX MEDLINE=97396144; PubMed=9245696;
RA Tanaka T., Tanaka M., Mitsui Y.;
RT "Analysis of expressed sequence tags (ESTs) of the parasitic protozoa
RT *Entamoeba histolytica*.";
RL Blochem. Biophys. Res. Commun. 236:611-615(1997).

```

DR      EMBL; AB002741; BAA21986.1; -.
DR      InterPro; IPR000004; SAPB.
DR      SMART; SM00118; SAPB; 1.
FT      NON_TER      1
FT      NON_TER      87
SO      SEQUENCE      87 AA; 9131 MW; FDDIC1B0FAB8B06E CRC64;

```

```
Query Match      49.28;  Score 63;  DB 5;  Length 87;
Best Local Similarity 54.58;  Pred. NO. 0.039;
Matches 12;  Conservative 3;  Mismatches 7;  Indels 0;  Gaps 0;
```

```
QY      1 GFATLCTKVLDFGIDKLIQLI 22
          | : ||||| : ||| || : |
Db      59 GLVATLCTPIASFGIAKLLASI 80
```

RESULT 4

ID	Accession	Source	Release Date	Version	Length	Score	Annotations	Keywords	Comments
049409	049409	PRELIMINARY	PRT	421	AA				
AC	049409								
DT	01-JUN-1998	(TREMBLrel. 06, Created)							
DT	01-JUN-1998	(TREMBLrel. 06, Last sequence update)							
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)							
DE	Protein kinase - like protein (Protein kinase-like protein).								
GN	FK5C5.120 OR AT4G18950.								
OS	Arabidopsis thaliana (Mouse-ear cress).								
OC	Eumaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;								
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;								
OC	eucotsids II; Brassicales; Brassicaceae; Arabidopsids.								
OX	NCBI_TaxID=3702;								
LN	[1]								
RP	SEQUENCE FROM N.A.								
RA	Bayram M., Pohl T., Welzenegger T., Bancroft I., Mewes H.W., Mayer K.								
RA	Schneeller C.								
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.								

RP SEQUENCE FROM N.A.
RA Pohl T., Weizenecker T., Mewes H.W., Lemcke K., Mayer K.
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AL161549; CAB78897.1; -
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000719; Euk_kinase.

DR Pfam; PF00023; ank; 3.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.

```
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00088; ANK_REPEAT; 2.
DR PROSITE; PS0297; ANK_REPEAT_REGION; 3.
```

KW ANK repeat; ATP-binding; Kinase; Repeat; Transferase.
 SO SEQUENCE 421 AA: 48173 MW: 6CC6FF878686A5AE CRC64:

Query Match	42.28;	Score 54;	DB 10;	Length 421;
Best Local Similarity	58.88;	Pred. NO. 4.6;		
Matches 10;	Conservative 3;	Mismatches 4;	Indels 0;	Gaps 0;

QY	9	KVLDFGIDKLIQLIEDK	25
		11111:11::111	
Db	258	KVADFGVSKLVTYKEDK	274

RESULT 5
Q93Z30
ID Q93Z30
PRELIMINARY: PRT: 459 AA

AC	Q32630;
DT	01-DEC-2001 (Tremblurel, 19, Created)
DT	01-DEC-2001 (Tremblurel, 19, Last sequence update)
DT	01-JUN-2002 (Tremblurel, 21, Last annotation update)

DE Arabidopsis thaliana (Mouse-ear cress).
 DE Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 OC Spermatophyta: Magnoliopsida: core eudicots: Rosidae
 OC Spermatophyta: Magnoliopsida: core eudicots: Rosidae

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
PB SEQUENCE FROM N A

RA Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
Kawaji T., Kim T.H., Kingston I.L., Kohn D., Kondo S., Kozlovskiy
L.A., Liang X., Lin W., Liu F., Liu J., Liu Z., Lu B., Maier G.,

RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurata
RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,


```

RESULT 6
09J3N7 ID 09J3N7 PRELIMINARY; PRT; 835 AA.
AC 09J3N7;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ORE51.
OS Human herpesvirus 3
CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10335;
RN [1]
RP SEQUENCE FROM N.A.
RA Araya T., Cohen J.I., Klutch M., Lektrom K., Yoshikawa T., Asano Y.,
RA Krause P.R.;
RT "Nucleotide sequences that distinguish Oka vaccine from parental Oka
RL and other varicella-zoster virus isolates."
RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
RL EMBL: AF206304; AAF61652.1; -.
DR InterPro: IPR003450; Herpes_orf_bp.
DR Pfam: PF02399; Herpes_orf_1p.
SO SEQUENCE 835 AA; 94383 MW; B0386877ACEA9FFB CRC64;

Query Match 41.4%; Score 53; DB 12; Length 835;
Best Local Similarity 43.5%; Pred. No. 13;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Oy 1 GPIATLCRKVLDFGIDKLIQLE 23
II II : III I::I:
Db 225 GFSGRTCTILRDMGIIDTLVRIRK 247

```

RX MEDLINE=98037514; PubMed=93714463.
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubots J.,
 RA Aldredge T., Bashitzaden R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumu W., Pothier B., Qu D.,
 RA Spadafora R., Viicare R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jivani N., Carruso A., Bush D., Seifer H., Patwell D., Prabhakar S.,
 RA McDowall S., Shlmer G., Goyal A., Pletrovski S., Church G.M.,
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of *Methanobacterium thermoautotrophicum*
 RT deltaH: functional analysis and comparative genomics";
 RJ J. Bacteriol. 179:7135-7155(1997).
 DR EMBL: AED00803; AAB84646.1; -.
 DR InterPro: IPR001450; 4Fe4S_ferredoxin.
 DR InterPro: IPR003741; DUF162.
 DR Pfam: PF02589; DUF162; 1.
 DR DR Pfam: PF00037; fer4; 2.
 DR PROSITE: PS00198; 4Fe4S_FERREDOXIN; 2.
 KW Iron-sulfur; Complete Proteome.
 SO SEQUENCE 396 AA; 42840 MW; 58004872C323A9BB CMC64;

Qy	3	1AFLCTKVLDEGLDQLQLIED	24
Db	207	LSLMDTFHMFEGVDKILVPOIED	228

RESULT	8
08TV25	
ID	08TV25
AC	08TV25;
DT	01-JUN-2002 (Tremblrel. 21, Created)
DT	01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT	01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE	Amino acid transporter.
GN	MK1576.
OS	Methanopyrus kandleri.
OC	Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC	Methanopyrus.
OX	NCBI_TaxID=2320;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=AV19 / DSM 6324 / JCM 9639;
RX	MEDLINE=21927647; PubMed=11930014;
RA	Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
RA	Shecherdinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA	Natale D.A., Rogozin I.B., Tatusev R.L., Wolf Y.I., Stetter K.O.,
RA	Mal'kh A.G., Koonin E.V., Kozayavkin S.A.;
RT	"The complete genome of hyperthermophilic <i>Methanopyrus kandleri</i> AV19
RT	and morphology of archaeal methanogens.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR	EMBL; AE010447; AA02789.1; -.
KW	Complete proteome.
Q0	SEQUENCE 491 AA; 52944 MW; 0A087E1EE2981B9B CRC64;

Query Match	39.8%	Score 51	DB 17	Length 491
Best Local Similarity	40.0%	Pred. No. 16		
Matches	10	Conservative	7	Mismatches 8
				Indels 0
				Gaps 0
Oy	1	GFATLCTKRVLDGIDKLIQLIEDK	25	-
		:	::	: :
Db	14	GFSTICCTVIGGYNVLTCLQVK	38	
RESULT 9				
Q8YNP6	ID	PRELIMINARY;	PRT;	520 AA.
Q8YNP6	ID	Q8YNP6;		
DT	01-MAR-2002	(TREMBLrel.	20.	Created)
DT	01-MAR-2002	(TREMBLrel.	20.	Last sequence update)
DT	01-JUN-2002	(TREMBLrel.	21.	Last annotation update)

DE Serine/threonine kinase.
 GN AL4518.
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria: Cyanobacteria: Nostocales: Nostocaceae: Nostoc.
 OX NCBI_TaxID=103690.
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Molk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriuchij M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.:
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).
 DR EMBL: AP003596; BAB76217.1; -
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_Thr_pkinase.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; TyrKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 KM Kinase: Complete proteome.
 SQ SEQUENCE 520 AA; 58905 MW; 2E6FE359B80ACF49 CRC64;

Query Match 39.8%; Score 51; DB 16; Length 520;
 Best Local Similarity 64.7%; Pred. No. 17;
 Matches 11; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 OY 9 KYLDKIDKLIQIIEBK 25
 DB 168 KYLDKIDKLIQIIEBK 184

RESULT 10
 09A314 PRELIMINARY: PRT: 213 AA.
 AC 09A314;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Hypothetical protein CC3392.
 GN CC3392.
 OS Caulobacter crescentus.
 OC Bacteria: Proteobacteria: alpha subdivision: Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647;
 RA Niseman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohba N., Maddock J.R.,
 RA Pollock I., Nelson W.C., Newton A.S., Stephens C., Phade D.D., Ely B.,
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
 RA Ullrich S.L., Venter J.C., Wolf A., Yamalhevyan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.:
 RT "Complete genome sequence of Caulobacter crescentus.";
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL: AE005999; AAK25354.1; -
 DR TIGR: CC3392; -
 KM Hypothetical protein: Complete proteome.
 SQ SEQUENCE 213 AA; 22524 MW; 6C99F2461AEDA2E CRC64;

Query Match 39.1%; Score 50; DB 16; Length 213;
 Best Local Similarity 40.0%; Pred. No. 9.7;
 Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

OY 1 GFATLCTKVLDFGIDKLIQIIEBK 25
 DB 177 GFATLCTKVLDFGIDKLIQIIEBK 201

RESULT 11
 09V112 PRELIMINARY: PRT: 233 AA.
 AC 09V112;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE CG14400 protein.
 GN CG14400.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:
 OC Eukaryota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha;
 OC Ephydroidea: Drosophilidae: Drosophila.
 OX NCBI_TaxID=7227;
 RN 11
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari P., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson R.A., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Relneert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spletter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
 RT "The genome sequence of Drosophila melanogaster.";
 RT Science 287:2185-2195(2000).
 DR EMBL: AE003668; AAF53937.1; -
 DR FLYbase: FBgn0032896; CG14400.
 SQ SEQUENCE 233 AA; 27298 MW; 41899178B71A1714 CRC64;

Query Match 39.1%; Score 50; DB 5; Length 233;
 Best Local Similarity 47.6%; Pred. No. 11;
 Matches 10; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
 OY 4 ATLCTKVLDFGIDKLIQIIEBK 24
 DB 5 ATLCTKVLDFGIDKLIQIIEBK 25

RESULT 12

09ELW9 PRELIMINARY: PRT: 816 AA.

AC 09ELW9: 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Origin binding protein.
 OS Cercopithecine herpesvirus 7.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=35245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.:
 RT "Complete Sequence of the Simian Varicella Virus Genome."
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF275348; AAG27226.1; -
 DR InterPro: IPR003593; AAA.ATPase.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR003450; Herpes_orf_bp.
 DR Pfam: PF02399; Herpes_orf_bp; 1.
 DR SMART: SM00382; AAA; 1.
 DR SMART: SM00487; DEXDC; 1.
 SO SEQUENCE 816 AA; 92834 MW; 6657D88021463CFF CRC64;

Query Match 39.1%; Score 50; DB 12; Length 816;
 Best Local Similarity 47.6%; Pred. No. 37;
 Matches 10; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 GFATICTKVLDFGIDKLIQL 21
 DB 214 GESSRCTFLADIGADALIEL 234

RESULT 13

008743 PRELIMINARY: PRT: 309 AA.

AC 008743: 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Chromosome XV reading frame ORF YOR292C.
 GN YOR292C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cziepluch C., Jauniaux J.C., Kordes E., Poirey R., Pujol A.,
 RA Tobiasch E.;
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MIPS:
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-97298310; PubMed-9153758;
 RX Poirey R., Cziepluch C., Tobiasch E., Pujol A., Kordes E.,
 RA Jauniaux J.C.;
 RT "Sequence and analysis of a 36.2 kb fragment from the right arm of
 RT yeast chromosome XV reveals 19 open reading frames including SNF2 (5'
 RT end), CPA1, STY41, a putative transport ATPase, a putative ribosomal
 RT protein and an SNF2 homologue."
 RL yeast 13:479-482(1997).
 DR EMBL; Z75200; CAA95520.1; -
 DR SGD; S0003818; YOR292C.
 SO SEQUENCE 309 AA; 36195 MW; A3BD32893F595536 CRC64;

Query Match 38.3%; Score 49; DB 3; Length 309;
 Best Local Similarity 56.2%; Pred. No. 20;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 4 ATLCTKVLDFGIDKLI 19
 DB 86 ATLCTNILLFGISDIL 101

RESULT 14

0943D4 PRELIMINARY: PRT: 629 AA.

AC 0943D4: 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE P0443D08.19 protein.
 GN P0443D08.19.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GAS) genomic DNA, chromosome 1, PAC
 RT clone:P0443D08."
 RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF003250; BAB64139.1; -
 DR InterPro: IPR000345; Cyt_c_heme_bind.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR005213; HgW.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF03578; HgW; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR Prodom: PD000001; Euk_pkinase; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
 DR ATP-binding; Transferase.
 SO SEQUENCE 629 AA; 70249 MW; A6D7E4F2753F2C83 CRC64;

Query Match 38.3%; Score 49; DB 10; Length 629;
 Best Local Similarity 56.2%; Pred. No. 41;
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 6 LCTKVLDFGIDKLIQL 21
 DB 504 LCPKIADFGIAGIKLHL 519

RESULT 15

096X52 PRELIMINARY: PRT: 551 AA.

AC 096X52: 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Phosphate transporter.
 GN PD185.
 OS Pholiota nameko.
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
 OC Agaricales; Strophariaceae; Pholiota.
 OX NCBI_TaxID=61267;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-N2;
 RA Tasaki Y., Kamiya Y., Hara T., Joh T.;
 RT "Isolation and sequencing of cDNA of genes induced by phosphate
 RT deficiency in cultured mycelia of Pholiota nameko."
 RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 DR EMBL: AB060641; BAB43910.1; -
 DR InterPro: IPR004738; Phos_permease.
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR TIGRfam: TIGR00887; 2A0109.1.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; UNKNOWN_1.
 KW Transmembrane.
 SQ SEQUENCE 551 AA: 59892 MW: 9FED9A5C7C89FFC1 CRC64;

Query Match 37.9%; Score 48.5; DB 3; Length 551;
 Best Local Similarity 50.0%; Pred. No. 43;
 Matches 13; Conservative 2; Mismatches 10; Indels 1; Gaps 1;

OY 1 GFIA-TLCTKVLDFGIDKLIQLIEDK 25
 ||||| : : |||||
 DB 402 GFIALTILFVILGFGYDKLISTPSSK 427

Search completed: May 9, 2003, 16:17:39
 Job time : 21.5294 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:57:44 ; Search time 40.95 Seconds
(without alignments)
29,286 Million cell updates/sec

Title: US-09-851-422B-1
Perfect score: 46
Sequence: 1 LVDRATCLR 9

Scoring table: BIOSUM62
Gap0 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneq/geneq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneq/geneq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneq/geneq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneq/geneq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneq/geneq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneq/geneq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneq/geneq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneq/geneq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneq/geneq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneq/geneq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneq/geneq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneq/geneq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneq/geneq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneq/geneq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneq/geneq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneq/geneq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneq/geneq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneq/geneq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneq/geneq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneq/geneq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneq/geneq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneq/geneq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	181	22 ABB36008	Peptide #3514 enco
2	46	100.0	182	22 AAB60289	Human Factor IX (h
3	46	100.0	412	16 AAB64266	Human Factor IX.
4	46	100.0	415	14 AAR35761	Factor IX (IX). H
5	46	100.0	454	16 AAR67710	Human Factor IX.
6	46	100.0	456	5 AAP40178	Part of the sequen
7	46	100.0	456	5 AAP40222	Sequence encoded b
8	46	100.0	461	6 AAP50311	Sequence of human
9	46	100.0	461	6 AAP50302	Sequence of human
10	46	100.0	461	6 AAP50019	Sequence of human

11	46	100.0	461	11	AAR05393	Mutant human facto
12	46	100.0	461	19	AAM40284	Human Factor IX pr
13	46	100.0	461	21	AAY97295	Human clotting fac
14	46	100.0	461	22	AAE10828	Human wild-type fa
15	46	100.0	461	22	AAB60281	Human factor IX (h
16	46	100.0	461	23	AAO21524	Protein of human f
17	46	100.0	461	23	AAM50568	Human Factor IX.
18	46	100.0	462	12	AAR10868	Recombinant human
19	89.1	415	20	AAY03203	Amino acid sequenc	
20	37	80.4	367	22	AAU47611	Proionibacterium
21	34	73.9	63	20	AAY11751	Human 5' EST secre
22	34	73.9	63	21	AAG00629	Human secreted pro
23	34	73.9	158	12	AAR11721	ENV93/HTLV-1-IIIB'
24	33	71.7	96	22	AAU14883	Novel bone marrow
25	33	71.7	101	21	AAY86427	Human gene 31-enco
26	33	71.7	138	20	AAY31395	HEV peptide z12-or
27	33	71.7	205	22	AAU47245	Proionibacterium
28	33	71.7	212	20	AAY28590	Human Factor D. H
29	33	71.7	250	11	AAR05421	Human adipisin/D en
30	33	71.7	250	11	AAR05772	Human adipisin gene
31	33	71.7	345	22	ABB63738	Drosophila melanog
32	33	71.7	575	22	AAM39439	Human polypeptide
33	33	71.7	581	22	ABB11366	Human KIAA0220 pro
34	33	71.7	877	22	AAM41225	Human polypeptide
35	33	71.7	877	22	ABB62624	Drosophila melanog
36	33	71.7	997	12	AAR15345	Mouse bullous pemp
37	33	71.7	1374	22	ABB01235	Novel human diapo
38	32	69.6	12	14	AAR30472	ICAM-3 peptide fra
39	32	69.6	38	22	AAM90306	Human immune/haema
40	32	69.6	83	22	ABB29363	Peptide #2014 enco
41	32	69.6	83	22	ABB34537	Peptide #1945 enco
42	32	69.6	83	22	ABB19946	Protein #1945 enco
43	32	69.6	83	22	AAM55323	Human brain expres
44	32	69.6	83	22	AAM67720	Human bone marrow
45	32	69.6	83	22	AAM15330	Peptide #1964 enco

ALIGNMENTS

RESULT 1
ABB36008
ID ABB36008 standard; Peptide: 181 AA.
XX
AC ABB36008;
DT 04-FEB-2002 (first entry)
XX
DE Peptide #3514 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.

FT /label- OTHER
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site
FT 19 /label- OTHER
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site
FT 20 /label- OTHER
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site
FT 25 /label- OTHER
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site
FT 26 /label- OTHER
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site
FT 29 /label- OTHER
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site
FT 32 /label- OTHER
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site
FT 35 /label- OTHER
FT /note= "gamma-carboxyglutamic acid"
FT Modified-site
FT 39 /label- OTHER
FT /note= "gamma-carboxyglutamic acid"
FT Disulfide-bond
FT 17..22 /label- OTHER
FT /note= "gamma-carboxyglutamic acid"
FT Disulfide-bond
FT 50..61 /label- OTHER
FT Disulfide-bond
FT 55..69 /label- OTHER
FT Disulfide-bond
FT 71..80 /label- OTHER
FT Disulfide-bond
FT 86..97 /label- OTHER
FT Disulfide-bond
FT 93..107 /label- OTHER
FT Disulfide-bond
FT 109..122 /label- OTHER
FT Disulfide-bond
FT 130..286 /label- OTHER
FT Disulfide-bond
FT 204..220 /label- OTHER
FT Disulfide-bond
FT 333..347 /label- OTHER
FT Disulfide-bond
FT 358..386 /label- OTHER
FT Region
FT 47..127 /label- OTHER
FT /note= "EGF-like region"
FT Region
FT 128..195 /label- OTHER
FT /note= "Factor-XIIa cleavage site"
FT Cleavage-site
FT 143..144 /label- OTHER
FT Modified-site
FT 155 /note= "PEG attachment site"
FT Modified-site
FT 165 /note= "PEG attachment site"
FT Cleavage-site
FT 178..179 /note= "Factor Xla cleavage site"
FT Domain
FT 196..412 /label- OTHER
FT /note= "Catalytic domain"
XX W09429370-A.
XX 22-DEC-1994.
XX PD
XX 07-JUN-1994; 94WO-US06388.
XX 08-JUN-1993; 93US-0073531.
XX (ENZO-) ENZON INC.
XX PA
XX Gilbert CW, Hallahan TW;
XX WPI; 1995-036423/05.
XX DR
XX New Factor IX-non-antigenic polymer conjugates - useful for
XX treatment of haemophilia B
XX PT
XX Disclousure; Fig. 1; 27pp; English.
XX PS
XX Human Factor-IX is a single chain glycoprotein having the sequence
CC given in AAR64266. Conjugates of Factor-IX with polymers such as

CC polyethylene glycol show reduced immunogenicity.
XX Sequence 412 AA:
SQ Query Match 100.0%; Score 46; DB 16; Length 412;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDRATCLR 9
Db 327 LVDRATCLR 335
RESULT 4
AAR35761
ID AAR35761 standard; protein: 415 AA.
XX AAR35761;
AC AAR35761;
DT 24-SEP-1993 (first entry)
DE Factor IX (IX).
XX PC; protein C; IX; Factor IX; X; Factor X; PT; prothrombin; VII;
XX Factor VII; CT; chymotrypsinogen; SP; serine protease; binding;
XX exosite; catalytic activity.
XX Homo sapiens.
OS
XX
FH Key
FH 1..145 Location/Qualifiers
FT Region
FT 146..180 /note= "Factor IX light chain"
FT Region
FT 181..415 /note= "Factor IX activation"
FT Region
FT /note= "Factor IX heavy chain"
FT Peptide
FT 395..409 /note= "exosite 1"
FT Peptide
FT 316..336 /note= "exosite 2"
FT Peptide
FT 316..330 /note= "pref. PC polypeptide; claim 2, page 136"
FT Peptide
FT 321..330 /note= "pref. PC polypeptide; claim 2, page 136"
FT Peptide
FT 395..409 /note= "pref. PC polypeptide; claim 2, page 136"
FT Peptide
FT 400..409 /note= "pref. PC polypeptide; claim 2, page 136"
FT Peptide
FT 400..414 /note= "pref. PC polypeptide; claim 2, page 136"
FT Peptide
FT 271..292 /note= "claim 6, page 137-138 describes an antibody
that reacts with Factor IX; fragments
316-330, 321-330, 395-409, 400-409 and
400-414 but not with fragment 271-292"
XX W09309804-A.
XX 27-MAY-1993.
XX PD
XX 18-NOV-1992; 92WO-US10242.
XX PF
XX 18-NOV-1991; 91US-0793989.
XX PR
XX (SCRI) SCRIPPS RES INST.
XX PA
XX GriffIn JH, Mesters RM;
XX WPI; 1993-182244/22.
XX DR
XX Serine protease derived-polypeptide(s) and anti-peptide
XX antibodies - for inhibiting coagulation and assaying for the
XX presence of serine protease in fluid samples

```
XX XX Disclosure: Page 126-128; 149pp; English.
PS PS
CC CC The PC polypeptides indicated in the Features Table inhibit
CC CC coagulation (they prevent binding of serine protease to natural
CC CC substrates), esp. when admin. to give an intravascular blood
CC CC concn. of 0.1-100 (pref. 0.5-10) microm.
CC CC NB: Sequences corresp. to SEQ ID NO 6, 7, 8 and 9 are described
CC CC in the specification but have not yet been added to the SEQUENCE
CC CC LISTING.
XX XX
SO Sequence 415 AA:
Query Match 100.0%; Score 46; DB 14; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDRATCLR 9
DB 330 LVDRATCLR 338
RESULT 5
AAR67710
ID AAR67710 standard; Protein: 454 AA.
XX XX
AC AAR67710;
XX XX
DT 20-JUL-1995 (first entry)
XX XX
DE Human Factor-IX.
XX XX
KW Factor-IX; blood-clotting; hemophilia B; gene therapy;
KW adenovirus; vector.
XX XX
OS Homo sapiens.
XX XX
PN WO9429471-A.
XX XX
PD 22-DEC-1994.
XX XX
PF 13-APR-1994; 94WO-US04075.
XX XX
PR 10-JUN-1993; 93US-0074920.
XX XX
PR 25-MAR-1994; 94US-0218335.
XX XX
PA (GENE-) GENETIC THERAPY INC.
XX XX
PI Connelly S, Kaleko M, Smith T;
XX XX
DR WPI: 1995-036495/05.
XX XX
DR N-PSDB: AA076017.
XX XX
PT New adenoviral vectors for treatment of haemophilia - contg. a
PT DNA sequence encoding a clotting factor, partic. Factor VIII or
PT Factor IX
XX XX
PS Disclosure: Fig. 11A-11B; 116pp; English.
XX XX
CC Human Factor-IX was used to construct recombinant adenovirus vectors
CC CC that produced therapeutic levels of the clotting factor when
CC CC administered to an animal host, potentially providing hemophilia B
CC CC gene therapy.
XX XX
SO Sequence 454 AA:
Query Match 100.0%; Score 46; DB 16; Length 454;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDRATCLR 9
DB 369 LVDRATCLR 377
```

```
RESULT 6
AAP40178
ID AAP40178 standard; Protein: 456 AA.
XX XX
AC AAP40178;
XX XX
DT 11-FEB-1992 (first entry)
XX XX
DE Part of the sequence of human factor IX.
XX XX
KW Christmas disease; therapy; haemophilia; factor IX; blood clotting;
KW diagnosis.
XX XX
OS Homo sapiens.
XX XX
FH Key Location/Qualifiers
FT Peptide 1..41
FT Protein 42..456
XX XX
PN WO8400560-A.
XX XX
PD 16-FEB-1994.
XX XX
PF 03-AUG-1983; 83WO-GB00191.
XX XX
PR 06-MAY-1983; 83GB-0012491.
PR 04-AUG-1982; 82GB-0022485.
XX XX
PA (NATR ) NATIONAL RES DEV CORP.
PA (BROW/) BROWNLEE G G.
XX XX
PI Brownlee G, Choo KH;
XX XX
DR WPI: 1984-049331/08.
XX XX
DR N-PSDB: AAN40142.
XX XX
PT Recombinant DNA cloning vehicles - useful in prodn. of factor IX
PT polypeptide and of diagnostic probes for Christmas disease
XX XX
PS Example: Fig 9; 71pp; English.
XX XX
CC The inventors claim DNA molecules comprising part or all of the
CC CC human factor IX DNA. The invention also includes cDNA derived from
CC CC human factor IX RNA. Specifically claimed are: recombinant DNA (the
CC CC phage present in clone lambda HIX-1) deposited as NCIB No. 11749;
CC CC Recombinant DNA in which the cloning vehicle is the modified pMT 153
CC CC plasmid present in E.coli NCIB No. 11747; Recombinant DNA in which
CC CC the bovine factor IX DNA sequence is contained in the recombinant
CC CC DNA transformed into E.coli to form a clone deposited as NCIB No.
CC CC 11748.
XX XX
SO Sequence 456 AA:
Query Match 100.0%; Score 46; DB 5; Length 456;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDRATCLR 9
DB 371 LVDRATCLR 379
RESULT 7
AAP40222
ID AAP40222 standard; Protein: 456 AA.
XX XX
AC AAP40222;
XX XX
DT 13-FEB-1992 (first entry)
XX XX
DE Sequence encoded by part of the sequence of human factor IX cDNA.
```



```

XX Haemophilia: Christmas disease; diagnosis; treatment.
KM Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX Peptide 1..41
XX Protein /label= signal
XX 42..456
XX GB2125409-A.
XX
XX 07-MAR-1984.
XX
XX 03-AUG-1983; 83GB-0020975.
XX
XX 16-MAY-1983; 83GB-0012490.
XX
XX 04-AUG-1982; 82GB-0022486.
XX
XX 03-AUG-1983; 83GB-0020975.
XX
XX (NATR ) NATIONAL RES DEV CORP.
XX
XX Brownlee GC, Choo KH;
XX
XX WPI; 1984-057898/10.
XX
XX N-PSDB; AAN40177.
XX
XX Prodn. of artificial human factor IX - by use of recombinant DNA
XX sequences for host transformation and cultivation
XX
XX Example; Fig 9: 49pp; English.
XX
XX The inventors claim a recombinant DNA having a human factor IX
XX sequence pref. at least 50 nucleotides long, esp. 75-27000
XX nucleotides. A cloning vector contg. foreign DNA is also claimed.
XX The foreign sequence pref. includes the whole of an exon sequence of
XX the human factor IX genome. The cloning vehicle may be a modified
XX PAT 153 plasmid. Also claimed is a labelled diagnostic probe
XX comprising a DNA molecule having a single- or double-stranded probe
XX sequence of 15 to 10000 nucleotides long Factor IX DNA sequence.
XX
XX SQ Sequence 456 AA:
XX
XX Query Match 100.0%; Score 46; DB 5; Length 456;
XX Best Local Similarity 100.0%; Pred. No. 1.1;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LVDRATCLR 9
XX
XX Db 371 LVDRATCLR 379
XX
XX RESULT 8
XX AAP50311
XX ID AAP50311 standard; Protein; 461 AA.
XX
XX AAP50311;
XX
XX 03-SEP-1991 (first entry)
XX
XX Sequence of human factor IX, encoded by DNA FIX.
XX
XX Blood clotting; haemophilia B.
XX
XX Homo sapiens.
XX
XX W08505125-A.
XX
XX 21-NOV-1985.
XX
XX 09-MAY-1984; 84WO-0007125.
XX
XX 09-MAY-1984; 84FR-0007125.
XX
XX

```

```

XX (TRAN-) TRANSGENE SA.
XX PA (JAVE/) JAVE M.
XX
XX
XX Jaye M, De La Salle H, Tolstoshev P, Lecocq JP;
XX
XX WPI; 1985-302926/48.
XX
XX N-PSDB; AAN50362.
XX
XX Cloning and expression vector for factor nine --for transforming
XX bacteria, yeast or mammalian cells
XX
XX Disclosure; Fig 3: 98pp; French.
XX
XX The inventors claim a vector which comprises a DNA SQ coding for
XX factor IX (FIX) and elements providing expression of this SQ in host
XX cells. The prepn. of DNA FIX is as follows. Polyadenylated mRNA is
XX isolated from human liver, used to make ss-DNA which is digested
XX with S1 and the fragments over 1 kb cloned in pBR 322. Clones are
XX selected by hybridisation with a 52-mer probe corresp. to the bovine
XX factor IX SQ (AAN50362).
XX
XX SQ Sequence 461 AA:
XX
XX Query Match 100.0%; Score 46; DB 6; Length 461;
XX Best Local Similarity 100.0%; Pred. No. 1.1;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 LVDRATCLR 9
XX
XX Db 376 LVDRATCLR 384
XX
XX RESULT 9
XX AAP50302
XX ID AAP50302 standard; Protein; 461 AA.
XX
XX AAP50302;
XX
XX 03-SEP-1991 (first entry)
XX
XX Sequence of human factor IX.
XX
XX Blood clotting; haemophilia B; vaccinia vector; cowpox vector.
XX
XX Homo sapiens.
XX
XX W08505376-A.
XX
XX 05-DEC-1985.
XX
XX 29-MAY-1985; 85WO-1104408.
XX
XX 05-OCT-1984; 84FR-0015294.
XX
XX 22-MAY-1984; 84FR-0007959.
XX
XX (TRAN-) TRANSGENE SA.
XX PA (DSAL/) DE LA SALLE H.
XX
XX De La Salle H, Drillion R, Altenburger W, Tolstoshev P, Lecocq JP;
XX
XX WPI; 1985-316884/50.
XX
XX N-PSDB; AAN50351.
XX
XX Vector for expressing factor 9 in vertebrate cells - contg.
XX poxvirus genome in which factor 9 gene is inserted
XX
XX Example; Fig 2: 59pp; French.
XX
XX The factor IX coding SQ (AAN50351) is first inserted into a non-
XX essential region of viral DNA, cloned in a bacterial plasmid. A
XX double-reciprocal recombination procedure is then used to transfer
XX the insert into the viral genome where it is propagated and

```

CC expressed.

XX Sequence 461 AA:

Query Match 100.0%; Score 46; DB 6; Length 461;
 Best Local Similarity 100.0%; Pred. NO. 1.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRATCLR 9
 |||||
 DB 376 LVDRATCLR 384

RESULT 10

AAP50019

ID AAP50019 standard; Protein: 461 AA.

XX AAP50019;

AC AAP50019;

DT 06-SEP-1991 (first entry)

XX Sequence of human factor IX.

XX Haemophilia B; blood clotting; factor IX.

XX Homo sapiens.

XX EPI62782-A.

XX 27-OCT-1985.

XX 21-MAY-1985; 85EP-0400999.

XX 05-OCT-1984; 84FR-0015294.

XX 22-MAY-1984; 84FR-0007959.

XX (TRAN-) TRANSGENE SA.

XX de la Salle H, Drillion R, Altenburger W, Tolloshev P;

XX Lécocq JP;

XX WPI: 1985-298122/48.

XX N-PSDB; AAN50049.

XX Vector for expressing factor 9 in vertebrate cells - comprises

XX pox-virus genome contg. factor 9 gene

XX Example: Fig 2; 56pp; French.

XX The inventors claim vectors for expressing factor IX, or an

XX analogous protein in vertebrate cells which consist of the genome of

XX poxvirus into which a gene coding for factor IX or an analogous

XX protein has been inserted. Cultivation of infected cells produces

XX very pure factor IX (a deficiency of which causes haemophilia B).

XX Sequence 461 AA:

XX Query Match 100.0%; Score 46; DB 6; Length 461;

XX Best Local Similarity 100.0%; Pred. NO. 1.1;

XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRATCLR 9
 |||||
 DB 376 LVDRATCLR 384

RESULT 11

ID AAR05393

XX AAR05393 standard; Protein: 461 AA.

XX AAR05393;

XX 12-OCT-1990 (first entry)

XX Mutant human factor IX (FIX).

XX Factor IX; haemophilia B; vaccina; ds.

XX Homo sapiens.

XX Key Location/Qualifier

XX mutation replace(44,Pro)

XX mutation replace(47,Tyr)

XX EPI373012-A.

XX 13-JUN-1990.

XX 07-NOV-1989; 89EP-0403065.

XX 09-NOV-1988; 88FR-0014635.

XX (TRAN-) TRANSGENE SA.

XX Meullien P;

XX WPI: 1990-180758/24.

XX N-PSDB; AAO05393.

XX DNA coding for human factor IX -

XX with mutation in pro coding sequence.

XX Disclosure; P; French.

XX The slightly altered gene product has a higher ratio of FIX

XX activity to FIX yield.

XX Sequence 461 AA:

XX Query Match 100.0%; Score 46; DB 11; Length 461;

XX Best Local Similarity 100.0%; Pred. NO. 1.1;

XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRATCLR 9
 |||||
 DB 376 LVDRATCLR 384

RESULT 12

AAM40284

ID AAM40284 standard; Protein: 461 AA.

XX AAM40284;

XX 16-JUN-1998 (first entry)

XX Human factor IX protease.

XX Factor X; factor IX; serine protease activity; catalytic domain; ZAD;

XX zymogen-activating domain; epidermal growth factor-like domain; EGF1;

XX EGF2; regulator; coagulation; fibrinolysis; homeostasis; X-ray structure;

XX detection; drug modelling; restriction protease.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Peptide 1..46

XX Protein /label= signal

XX Domain /label= Factor IX

XX Domain /label= EGF2

XX Domain /label= catalytic domain

XX WO9747737-A1.


```

XX 27-SEP-2001.
PD 21-MAR-2001: 2001WO-EP03220.
XX
XX 22-MAR-2000: 2000EP-0106225.
PR 08-MAY-2000: 2000US-203249P.
XX
XX (OCTA-) OCTAGENE GMBH.
XX
XX Hauser C, Hoerster A, Schroeder C, Lehnerer M.
XX
XX MPI: 2001-590175/66.
XX
XX N-PSDB: AAD18173.
XX
XX Recombinantly producing human blood coagulation factors VIII and IX for
PT use in treating hemophilia -
XX
XX Claim 21: Page 73-74, 104pp: English.
XX
XX The present invention relates to an improved method for the production
CC of recombinant human blood clotting factors, especially factors VIII and
CC IX, using an immortalised cell line stably expressing viral transcription
CC activators and carrying a vector comprising a promoter and a sequence
CC encoding the blood coagulation factor. The factor VIII multien or a gene
CC transfer vector is used in the preparation of agents for treating
CC haemophilia, especially haemophilia A. The present sequence is human
CC wild-type factor IX protein encoded by vector PTGf36hyg.
XX
XX Sequence 461 AA:
SQ
Query Match 100.0%; Score 46; DB 22; Length 461;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
DB 376 LVDRATCLR 384
RESULT 15
AAB60281
ID AAB60281 standard: Protein: 461 AA.
XX
XX AAB60281:
AC
XX
XX 30-MAR-2001 (first entry)
DT
XX
XX Human factor IX (hFIX) protein.
DE
XX
XX Age-related gene regulation: liver-specific; gene expression;
XX human factor IX; hFIX; AE5'; AE3'; age-regulatable expression construct;
XX antisense therapy; gene therapy; thrombosis; cardiovascular disease;
XX diabetes; Alzheimer's disease; Parkinson's disease; cancer; osteoporosis;
XX osteoarthritis; dementia.
XX
XX Homo sapiens.
OS
XX
XX WO200075279-A2.
PN
XX
XX 14-DEC-2000.
PD
XX
XX 06-JUN-2000; 2000WO-US15728.
PF
XX
XX 09-JUN-1999; 99US-0328925.
PR
XX
XX (UNMI ) UNITV MICHIGAN.
PA
XX
XX Kurachi K, Kurachi S;
PI
XX
XX MPI: 2001-061708/07.
DR
XX
XX N-PSDB: AAF54018.
XX

```

```

PT New regulatory elements that control age-related gene expression,
PT useful in gene therapy and for reducing Factor IX expression -
XX
XX Disclosure: Fig 8A-E, 225pp: English.
XX
XX The invention relates to nucleic acid sequences which regulate gene
CC expression in an age-related manner and/or in a liver-specific manner.
CC The invention identifies regions of the human factor IX (hFIX) gene, and
CC a region of the human protein C (hPC) gene, which are age-related
CC regulatory sequences. The hFIX age-related regulatory sequences are
CC designated AE5' (AAF54016) and AE3' (AAF54017) and are found in the 5'
CC UTR (at position 2164-2165 of AAF54018) and 3' UTR (at position
CC 34383-35653 of AAF54018) respectively. These elements act synergistically
CC to increase hFIX levels over the lifespan of an individual; however, they
CC can independently exert effects on hFIX mRNA in an age-related manner.
CC with AE5' acting to stabilise hFIX mRNA, and AE3' acting to increase hFIX
CC mRNA levels, over time. AE5' also directs liver-specific expression. The
CC hPC gene age-related regulatory sequence is found in the 5' UTR
CC (AAF54081), and contains two PEA-3 (polyoma virus activator 3) elements
CC 5'-GAGGAAA-3' and 5'-GAGGAG-3'. The age-related regulatory sequences of
CC the invention, along with their homologues, variants and fragments, may
CC be used in the construction of recombinant expression vectors for the
CC expression of a desired sequence in an age-related fashion in a host
CC cell. Preferred target genes for expression in such age-regulatable
CC expression vectors include those encoding proteins involved in blood
CC coagulation (e.g., the pro-coagulants factor IX and factor VIII, and the
CC anti-coagulants protein C and antithrombin III), human
CC alpha-1-antitrypsin, PEA-3 protein and reporter proteins such as
CC luciferase. Preferred promoters for use in such age-regulatable
CC expression vectors include the human factor IX promoter, the T7 promoter,
CC the T3 promoter and the SP6 promoter. The expression vectors of the
CC invention may be used in gene therapy to provide age-related and/or
CC liver-specific expression of target genes. Age-regulatable constructs may
CC be used in the treatment of such age-related conditions such as
CC thrombosis, cardiovascular disease, diabetes, Alzheimer's disease,
CC Parkinson's disease, cancer, osteoporosis, osteoarthritis and dementia.
CC Specifically, they may be used to express factor IX antisense mRNA in the
CC treatment of thrombotic conditions associated with the natural
CC age-related rise in factor IX expression. Transgenic cells or animals
CC that contain vectors of the invention are useful as models of these
CC diseases, in screening for potential therapeutic agents and for studying
CC normal processes such as ageing and gene expression. Fragments and
CC homologues of age-related regulatory sequences, are useful as probes to
CC detect, isolate or identify other such sequences in samples. The present
XX sequence represents hFIX.
XX
XX Sequence 461 AA:
SQ
Query Match 100.0%; Score 46; DB 22; Length 461;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
DB 376 LVDRATCLR 384

```

Search completed: May 6, 2003, 15:01:01
Job time : 42.95 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 14:59:30 ; Search time 14.4 Seconds
(without alignments)
18.389 Million cell updates/sec

Title: US-09-851-422B-1

Perfect score: 46
Sequence: 1 LVDRATCLR 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A-COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B-COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	100.0	40	2	US-08-797-842-11 Sequence 11, Appl
2	46	100.0	40	2	US-08-797-842-12 Sequence 12, Appl
3	46	100.0	235	4	US-08-944-483-48 Sequence 48, Appl
4	46	100.0	415	1	US-08-073-531B-1 Sequence 1, Appl
5	46	100.0	415	1	US-08-295-411-2 Sequence 2, Appl
6	46	100.0	415	2	US-08-955-471-2 Sequence 2, Appl
7	46	100.0	415	2	US-08-766-288-1 Sequence 1, Appl
8	46	100.0	415	5	PCT-US92-10242-2 Sequence 2, Appl
9	46	100.0	461	3	US-08-742-877-2 Sequence 2, Appl
10	46	100.0	461	4	US-09-053-871A-21 Sequence 21, Appl
11	46	100.0	461	6	US521070-2 Patent No. 5521070
12	33	71.7	228	4	US-08-944-483-44 Sequence 44, Appl
13	33	71.7	228	6	5223425-5 Patent No. 5223425
14	33	71.7	220	6	5223425-8 Patent No. 5223425
15	33	71.7	253	6	5223425-4 Patent No. 5223425
16	32	69.6	239	4	US-08-944-483-61 Sequence 61, Appl
17	32	69.6	268	1	US-08-270-584A-2 Sequence 2, Appl
18	32	69.6	268	1	US-08-568-031-2 Sequence 2, Appl
19	32	69.6	268	2	US-08-765-192-2 Sequence 2, Appl
20	32	69.6	268	2	US-08-966-319-2 Sequence 2, Appl
21	32	69.6	268	3	US-09-153-304-2 Sequence 2, Appl
22	32	69.6	268	3	US-09-199-793-2 Sequence 2, Appl
23	32	69.6	434	2	US-08-487-113D-120 Sequence 120, App
24	32	69.6	434	2	US-08-720-420A-120 Sequence 120, App
25	32	69.6	518	4	US-09-240-915-8 Sequence 8, Appl
26	32	69.6	518	4	US-09-591-435-8 Sequence 8, Appl
27	32	69.6	547	1	US-08-314-615-1 Sequence 1, Appl

28	32	69.6	547	1	US-08-314-362-1 Sequence 1, Appl
29	32	69.6	547	1	US-08-473-981A-6 Sequence 6, Appl
30	32	69.6	547	1	US-08-433-010-1 Sequence 1, Appl
31	32	69.6	547	1	US-08-482-882-1 Sequence 1, Appl
32	32	69.6	547	2	US-08-483-389-1 Sequence 1, Appl
33	32	69.6	547	2	US-08-487-113D-1 Sequence 1, Appl
34	32	69.6	547	2	US-08-473-503-1 Sequence 1, Appl
35	32	69.6	547	2	US-08-483-932-1 Sequence 1, Appl
36	32	69.6	547	2	US-08-474-087-6 Sequence 6, Appl
37	32	69.6	547	2	US-08-720-420A-1 Sequence 1, Appl
38	32	69.6	547	3	US-08-714-017-1 Sequence 1, Appl
39	32	69.6	547	3	US-08-863-790-1 Sequence 1, Appl
40	32	69.6	547	3	US-08-475-680-1 Sequence 1, Appl
41	32	69.6	547	4	US-08-296-749-1 Sequence 1, Appl
42	31	67.4	77	4	US-09-134-001C-3500 Sequence 3500, Ap
43	31	67.4	222	6	5223425-6 Patent No. 5223425
44	31	67.4	223	4	US-09-516-914-13 Sequence 13, Appl
45	31	67.4	846	1	US-08-149-103-3 Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-797-842-11
Sequence 11, Application US/08797842
Patent No. 5932706
GENERAL INFORMATION:
APPLICANT: Meritens, Koenraad et al
TITLE OF INVENTION: Antibodies specific for a haemostatic protein,
TITLE OF INVENTION: their use for isolating intact protein, haemostatic composi,
TITLE OF INVENTION: of proteolytic cleavage products of the protein
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Michaelson and Wallace
STREET: Parkway 109 Office Center, 328 Newman Springs
CITY: Red Bank
STATE: New Jersey
COUNTRY: USA
ZIP: 07701
COMPUTER READABLE FORM:
MEDIUM TYPE: 3 1/2" 1.44 Mbyte IBM compatible diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS-DOS, Windows for Workgroups 3.11
SOFTWARE: Microsoft word for Windows 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/797,842
FILING DATE: 10-Feb-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/381,891
FILING DATE: February 8, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Michaelson, Peter L.
REGISTRATION NUMBER: 30090
REFERENCE/DOCKET NUMBER: Stitching-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)530-6671
TELEFAX: (908)530-6584
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
US-08-797-842-11-
Query Match 100.0%; Score 46; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Caps 0;

OY 1 LVDRATCLR 9
111111111
DB 23 LVDRATCLR 31

RESULT 2

US-08-797-842-12
Sequence 12, Application US/08797842

Patent No. 5932706

GENERAL INFORMATION:

APPLICANT: Mertens, Koenraad et al

TITLE OF INVENTION: Antibodies specific for a haemostatic protein,

TITLE OF INVENTION: their use for isolating intact protein, haemostatic composite

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Michaelson and Wallace

STREET: Parkway 109 Office Center, 328 Newman Springs

STREET: Road, P. O. Box 8489

CITY: Red Bank

STATE: New Jersey

COUNTRY: USA

ZIP: 07701

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2" 1.44 MByte IBM compatible diskette

COMPUTER: IBM PC

OPERATING SYSTEM: MS-DOS, Windows for Workgroups 3.11

SOFTWARE: Microsoft Word for Windows 6

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/797,842

FILING DATE: 10-Feb-1997

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/381,891

FILING DATE: February 8, 1995

ATTORNEY/AGENT INFORMATION:

NAME: Michaelson, Peter L.

REGISTRATION NUMBER: 30090

REFERENCE/DOCKET NUMBER: Stitclung-5

TELECOMMUNICATION INFORMATION:

TELEPHONE: (908)530-6671

TELEFAX: (908)530-6584

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 40 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: peptide

HYPOTHEICAL: NO

US-08-797-842-12

Query Match 100.0%; Score 46; DB 2; Length 40;

Best Local Similarity 100.0%; Pred. No. 0.019; 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRATCLR 9
111111111

DB 12 LVDRATCLR 20

RESULT 3

US-08-944-483-48

Sequence 48, Application US/08944483

Patent No. 6232456

GENERAL INFORMATION:

APPLICANT: COHEN, MAURICE

APPLICANT: COLPITS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GRANADOS, EDWARD N.

APPLICANT: KLASIS, MICHAEL R.

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STEWART, KENT D.
APPLICANT: STROUPE, STEVEN D.
TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
TITLE OF INVENTION: OF THE PROSTATE
NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/944,483

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Becker, Cheryl L.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6183.US.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 235 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: NO. 6232456

US-08-944-483-48

Query Match 100.0%; Score 46; DB 4; Length 235;

Best Local Similarity 100.0%; Pred. No. 0.12; 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRATCLR 9
111111111

DB 150 LVDRATCLR 158

RESULT 4

US-08-073-531B-1

Sequence 1, Application US/08073531B

Patent No. 5621039

GENERAL INFORMATION:

APPLICANT: Hallahan, et al.

TITLE OF INVENTION: Factor IX - polymeric conjugates

NUMBER OF SEQUENCES: 1

CORRESPONDENCE ADDRESS:

ADDRESSEE: GALGANO & BURKE

STREET: 300 Radpro Drive

CITY: Hauppauge

STATE: New York

COUNTRY: USA

ZIP: 11788

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb Storage

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS 6.0

SOFTWARE: WordPerfect 6.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/073,531B

FILING DATE: June 8, 1993
CLASSIFICATION: 525
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: GALIANO & BURKE
REGISTRATION NUMBER: 30,735
REFERENCE/DOCKET NUMBER: 128-7
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 582-6161
TELEFAX: (516) 582-6191
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 Amino Acids
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown to applicant
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
ORIGINAL SOURCE:
ORGANISM:
INDIVIDUAL ISOLATE:
CELL TYPE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
RELEVANT RESIDUES IN SEQ ID NO: 1. FROM 1 TO 415.
US-08-073-531B-1
Query Match 100.0%; Score 46; DB 1; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDRATCLR 9
DB 330 LVDRATCLR 338
RESULT 5
US-08-295-411-2
Sequence 2, Application US/08295411
Patent No. 5679639
GENERAL INFORMATION:
APPLICANT: Griffin, John H.
APPLICANT: Westers, Rolf M.
TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Office of Patent Counsel, The Scripps
ADDRESS: Research Institute
STREET: 10666 No. 5679639th Torrey Pines Road, TPC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/295,411
FILING DATE: 22-AUG-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,989
FILING DATE: 18-NOV-1991
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSP1263.0C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Region
LOCATION: 1..145
OTHER INFORMATION: /note= "Factor IX Light Chain"
FEATURE:
NAME/KEY: Region
LOCATION: 146..180
OTHER INFORMATION: /note= "Factor IX Activation
OTHER INFORMATION: Peptide"
FEATURE:
NAME/KEY: Region
LOCATION: 181..415
OTHER INFORMATION: /note= "Factor IX Heavy Chain"
US-08-295-411-2
Query Match 100.0%; Score 46; DB 1; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LVDRATCLR 9
DB 330 LVDRATCLR 338
RESULT 6
US-08-955-471-2
Sequence 2, Application US/08955471
Patent No. 5968751
GENERAL INFORMATION:
APPLICANT: Griffin, John H.
APPLICANT: Westers, Rolf M.
TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Office of Patent Counsel, The Scripps
ADDRESS: Research Institute
STREET: 10666 No. 5968751th Torrey Pines Road, TPC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,471
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/295,411
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Filling, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSN1263.0C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Region
LOCATION: 1..145
OTHER INFORMATION: /note="Factor IX Light Chain"
FEATURE:
NAME/KEY: Region
LOCATION: 146..180
OTHER INFORMATION: /note="Factor IX Activation"
OTHER INFORMATION: Peptide"
FEATURE:
NAME/KEY: Region
LOCATION: 181..415
OTHER INFORMATION: /note="Factor IX Heavy Chain"
US-08-955-471-2
Query Match 100.0%; Score 46; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. NO. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
|||||
DB 330 LVDRATCLR 338
RESULT 7
US-08-766-288-1
Sequence 1, Application US/08766288
Patent No. 5969040
GENERAL INFORMATION:
APPLICANT: Hallahan, et al.
TITLE OF INVENTION: Factor IX - Polymeric Conjugates
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: GALGANO & BURKE
STREET: 300 Rabro Drive
CITY: Hauppauge
STATE: New York
COUNTRY: USA
ZIP: 11788
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb Storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS 6.0
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/766,288
FILING DATE:
CLASSIFICATION: 525
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/073,531
FILING DATE: June 8, 1993
ATTORNEY/AGENT INFORMATION:
NAME: GALGANO & BURKE

REGISTRATION NUMBER: 30,735
REFERENCE/DOCKET NUMBER: 128-7 (DIV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 582-6161
TELEFAX: (516) 582-6191
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 Amino Acids
TYPE: Amino Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown to applicant
MOLECULE TYPE:
HYPOTHETICAL:
ANTI-SENSE:
ORIGINAL SOURCE:
ORGANISM:
INDIVIDUAL ISOLATE:
CELL TYPE:
IMMEDIATE SOURCE:
LIBRARY:
CLONE:
PUBLICATION INFORMATION:
AUTHORS:
TITLE:
TITLE: JOURNAL:
VOLUME:
ISSUE:
PAGES:
DATE:
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 415.
US-08-766-288-1
Query Match 100.0%; Score 46; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. NO. 0.21;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LVDRATCLR 9
|||||
DB 330 LVDRATCLR 338
RESULT 8
PCT-US92-10242-2
Sequence 2, Application PC/TUS9210242
GENERAL INFORMATION:
APPLICANT: Griffin, John H.
TITLE OF INVENTION: Serine Protease-Derived Polypeptides and
TITLE OF INVENTION: Anti-Peptide Antibodies, Systems and Therapeutic Methods
TITLE OF INVENTION: for Inhibiting Coagulation
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Office of Patent Counsel, The Scripps
ADDRESS: Research Institute
STREET: 10666 North Torrey Pines Road, TPC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10242
FILING DATE: 19921118
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/793,989
FILING DATE: 18-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Filling, Thomas

REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRO472P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 415 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: Region
LOCATION: 1..145
OTHER INFORMATION: /note="Factor IX Light Chain"
FEATURE:
NAME/KEY: Region
LOCATION: 146..180
OTHER INFORMATION: /note="Factor IX Activation"
OTHER INFORMATION: peptide
FEATURE:
NAME/KEY: Region
LOCATION: 181..415
OTHER INFORMATION: /note="Factor IX Heavy Chain"
PCT-US92-10242-2

Query Match 100.0%; Score 46; DB 5; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRATCLR 9
|||||
DB 330 LVDRATCLR 338

RESULT 9
US-08-742-877-2
Sequence 2, Application US/08742877
Patent No. 6046380
GENERAL INFORMATION:
APPLICANT: CLARK, Anthony J.
TITLE OF INVENTION: DNA SEQUENCES
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/742,877
FILING DATE: 01-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9408717.8
FILING DATE: 03-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: FLESHNER, RAZ E.
REGISTRATION NUMBER: 34,331
REFERENCE/DOCKET NUMBER: 0623.0470001/REF
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 461 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-742-877-2

Query Match 100.0%; Score 46; DB 3; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRATCLR 9
|||||
DB 376 LVDRATCLR 384

RESULT 10
US-09-053-871A-21
Sequence 21, Application US/09053871A
Patent No. 6315995
GENERAL INFORMATION:
APPLICANT: Plinsky, David J.
APPLICANT: Stern, David
APPLICANT: Rose, Eric
APPLICANT: Solomon, Robert A.
APPLICANT: Schmidt, Ann Marie
TITLE OF INVENTION: METHODS FOR TREATING AN ISCHEMIC DISORDER AND IMPROVING
FILE REFERENCE: 51917-B
CURRENT APPLICATION NUMBER: US/09/053,871A
CURRENT FILING DATE: 1998-04-01
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 21
LENGTH: 461
TYPE: PRT
ORGANISM: Homo Sapien
US-09-053-871A-21

Query Match 100.0%; Score 46; DB 4; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRATCLR 9
|||||
DB 376 LVDRATCLR 384

RESULT 11
US-09-053-871A-21
Sequence 21, Application US/09053871A
Patent No. 6315995
GENERAL INFORMATION:
APPLICANT: Plinsky, David J.
APPLICANT: Stern, David
APPLICANT: Rose, Eric
APPLICANT: Solomon, Robert A.
APPLICANT: Schmidt, Ann Marie
TITLE OF INVENTION: METHODS FOR TREATING AN ISCHEMIC DISORDER AND IMPROVING
FILE REFERENCE: 51917-B
CURRENT APPLICATION NUMBER: US/09/053,871A
CURRENT FILING DATE: 1998-04-01
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 21
LENGTH: 461
TYPE: PRT
ORGANISM: Homo Sapien
US-09-053-871A-21

Query Match 100.0%; Score 46; DB 6; Length 461;
Best Local Similarity 100.0%; Pred. No. 0.23;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LVDRATCLR 9
|||||

DB 376 LVDRATCLR 384

RESULT 12

US-08-944-483-44

Sequence 44, Application US/08944483
Patent No. 6232456

GENERAL INFORMATION:

APPLICANT: COHEN, MAURICE

APPLICANT: COLPITTS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GRANADOS, EDWARD N.

APPLICANT: KLAS, MICHAEL R.

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STEWART, KENT D.

APPLICANT: STROUPE, STEVEN D.

TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS

TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES

NUMBER OF SEQUENCES: 76

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/944,483

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Becker, Cheryl L.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6183-US.01

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 228 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6232456c

US-08-944-483-44

Query Match 71.7% Score 33; DB 4; Length 228;
Best Local Similarity 66.7% Pred. No. 36;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9

DB 148 VLDRATCLR 156

RESULT 13

5223425-5

Patent No. 5223425

APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,

BARRY M.; WHITE, TYLER R.

TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

D ACTIVITY

NUMBER OF SEQUENCES: 19

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/277,963

FILING DATE: 30-NOV-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 34,203

FILING DATE: 02-APR-1987

SEQ ID NO: 5;

LENGTH: 238

Query Match 71.7% Score 33; DB 6; Length 238;
Best Local Similarity 66.7% Pred. No. 36;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9

DB 158 VLDRATCLR 166

RESULT 14

5223425-4

Patent No. 5223425

APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,

BARRY M.; WHITE, TYLER R.

TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

D ACTIVITY

NUMBER OF SEQUENCES: 19

CURRENT APPLICATION DATA:

FILING DATE: 30-NOV-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 34,203

FILING DATE: 02-APR-1987

SEQ ID NO: 4;

LENGTH: 250

Query Match 71.7% Score 33; DB 6; Length 250;
Best Local Similarity 66.7% Pred. No. 40;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9

DB 170 VLDRATCLR 178

RESULT 15

5223425-8

Patent No. 5223425

APPLICANT: FLIER, JEFFREY S.; SPIEGELMAN, BRUCE M.; ROSEN,

BARRY M.; WHITE, TYLER R.

TITLE OF INVENTION: DNA ENCODING HUMAN ADIPSIN WITH COMPLEMENT

D ACTIVITY

NUMBER OF SEQUENCES: 19

CURRENT APPLICATION DATA:

FILING DATE: 30-NOV-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 34,203

FILING DATE: 02-APR-1987

SEQ ID NO: 8;

LENGTH: 253

Query Match 71.7% Score 33; DB 6; Length 253;
Best Local Similarity 66.7% Pred. No. 40;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9

DB 173 VLDRATCLR 181

Mon May 12 15:29:54 2003

us-09-851-422b-1.ra1

Page 7

Search completed: May 6, 2003, 15:04:05
Job time : 16.4 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 6, 2003, 15:01:10 ; Search time 18 Seconds
(without alignments)
43.143 Million cell updates/sec

Title: US-09-851-422B-1

Perfect score: 46
Sequence: 1 LVDRATCLR 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published_Applications_AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	100.0	181	US-09-864-761-46600	Sequence 46600, A
2	46	100.0	461	US-10-132-829-5	Sequence 5, Appl1
3	46	100.0	461	US-09-884-901-3	Sequence 3, Appl1
4	41	89.1	415	US-09-118-748-2	Sequence 2, Appl1
5	33	71.7	101	US-10-012-542-351	Sequence 351, App
6	33	71.7	228	US-09-821-255-2	Sequence 2, Appl1
7	32	69.6	44	US-10-001-843-188	Sequence 188, App
8	32	69.6	83	US-09-864-761-35244	Sequence 35244, A
9	32	69.6	252	US-10-117-323-38	Sequence 38, Appl1
10	32	69.6	434	US-09-753-436-150	Sequence 120, App
11	32	69.6	547	US-10-047-542-65	Sequence 65, Appl1
12	32	69.6	547	US-09-753-436-1	Sequence 1, Appl1
13	32	69.6	2125	US-09-919-172-29	Sequence 29, Appl1
14	31	67.4	24	US-10-012-542-362	Sequence 362, App
15	31	67.4	223	US-09-870-162A-13	Sequence 13, Appl1
16	31	67.4	338	US-09-886-055-205	Sequence 205, App
17	31	67.4	382	US-09-731-872-312	Sequence 312, App
18	31	67.4	429	US-10-012-542-366	Sequence 366, App
19	31	67.4	466	US-09-731-872-311	Sequence 311, App

20	31	67.4	470	9	US-10-012-542-365	Sequence 365, App
21	31	67.4	873	9	US-10-167-264-2	Sequence 2, Appl1
22	30	65.2	10	9	US-09-572-404B-1740	Sequence 1740, App
23	30	65.2	215	10	US-09-925-297-660	Sequence 660, App
24	30	65.2	363	9	US-10-045-815-2	Sequence 2, Appl1
25	30	65.2	374	9	US-10-255-817-10	Sequence 10, Appl1
26	27	63.0	736	9	US-10-045-815-4	Sequence 4, Appl1
27	27	63.0	80	9	US-10-092-154-626	Sequence 626, App
28	29	63.0	80	10	US-09-764-847-626	Sequence 626, App
29	29	63.0	193	10	US-09-864-761-34304	Sequence 34304, A
30	29	63.0	282	9	US-10-167-555-2	Sequence 2, Appl1
31	29	63.0	300	9	US-10-224-414-2	Sequence 2, Appl1
32	29	63.0	326	9	US-10-108-915-46	Sequence 46, Appl1
33	29	63.0	459	9	US-09-759-595-3	Sequence 3, Appl1
34	29	63.0	460	9	US-10-001-254-16	Sequence 16, Appl1
35	29	63.0	460	9	US-10-001-254-28	Sequence 28, Appl1
36	29	63.0	460	9	US-09-759-595-1	Sequence 1, Appl1
37	29	63.0	909	9	US-09-759-595-1	Sequence 190, App
38	29	63.0	909	12	US-10-078-929-168	Sequence 168, App
39	28	60.9	10	9	US-09-572-404B-1749	Sequence 1749, App
40	28	60.9	35	9	US-10-083-357-1143	Sequence 1143, App
41	28	60.9	57	10	US-09-864-761-36792	Sequence 36792, A
42	28	60.9	70	12	US-10-107-371-1	Sequence 1, Appl1
43	28	60.9	75	9	US-10-104-755-12	Sequence 12, Appl1
44	28	60.9	77	9	US-09-792-793A-14	Sequence 14, Appl1
45	28	60.9	119	9	US-10-141-965-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-864-761-46600
Sequence 46600, Application US/09864761
Patent No. US2002048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30

```
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO: 46600
;; LENGTH: 181
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL033403.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.5
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
;; OTHER INFORMATION: EST_HUMAN HIT: 128608.1, EVALUATE 3.00e-36
;; OTHER INFORMATION: SWISSPROT HIT: P00740, EVALUATE 1.00e-107
US-09-864-761-46600
```

```
Query Match
Best Local Similarity 100.0%; Score 46; DB 10; Length 181;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LVDRATCLR 9
|||||
Db 96 LVDRATCLR 104
```

```
RESULT 2
US-10-132-829-5
;; Sequence 5, Application US/10132829
;; Publication No. US20030044982A1
;; GENERAL INFORMATION:
;; APPLICANT: Chien, Kenneth R
;; APPLICANT: Hoshijima, Masahiko
;; TITLE OF INVENTION: Method to treat hemophilia by hepatic gene transfer of Factor VII
;; FILE REFERENCE: 6627-PA1170
;; CURRENT APPLICATION NUMBER: US/10/132,829
;; CURRENT FILING DATE: 2002-04-25
;; PRIOR APPLICATION NUMBER: 60/286,314
;; PRIOR FILING DATE: 2001-04-25
;; NUMBER OF SEQ ID NOS: 5
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO: 5
;; LENGTH: 461
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-132-829-5
```

```
Query Match
Best Local Similarity 100.0%; Score 46; DB 9; Length 461;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LVDRATCLR 9
|||||
Db 376 LVDRATCLR 384
```

```
RESULT 3
US-09-884-901-3
;; Sequence 3, Application US/09884901
;; Patent No. US20020076798A1
;; GENERAL INFORMATION:
;; APPLICANT: Miao, Carol
;; APPLICANT: Kay, Mark
;; TITLE OF INVENTION: Liver-Specific Gene Expression Cassettes, and Methods of Use
;; FILE REFERENCE: USFW-1-17396
;; CURRENT APPLICATION NUMBER: US/09/884,901
;; CURRENT FILING DATE: 2001-06-18
```

```
;; PRIOR APPLICATION NUMBER: US 60/212,902
;; PRIOR FILING DATE: 2000-06-20
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: PatentIn version 3.0
;; SEQ ID NO: 3
;; LENGTH: 461
;; TYPE: PRT
;; ORGANISM: Homo sapien
US-09-884-901-3
```

```
Query Match
Best Local Similarity 100.0%; Score 46; DB 10; Length 461;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LVDRATCLR 9
|||||
Db 376 LVDRATCLR 384
```

```
RESULT 4
US-09-118-748-2
;; Sequence 2, Application US/09118748A
;; Patent No. US20020031799A1
;; GENERAL INFORMATION:
;; APPLICANT: Stafford, Darrel W.
;; APPLICANT: Chang, Jinli
;; TITLE OF INVENTION: Factor IX Antihemophilic Factor with Increased Clotting
;; FILE REFERENCE: 5470-183
;; CURRENT APPLICATION NUMBER: US/09/118,748A
;; CURRENT FILING DATE: 1998-07-17
;; EARLIER APPLICATION NUMBER: 60/053,571
;; EARLIER FILING DATE: 1997-07-21
;; NUMBER OF SEQ ID NOS: 2
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO: 2
;; LENGTH: 415
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-118-748-2
```

```
Query Match
Best Local Similarity 89.1%; Score 41; DB 10; Length 415;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LVDRATCLR 8
|||||
Db 330 LVDRATCLR 337
```

```
RESULT 5
US-10-012-542-351
;; Sequence 351, Application US/10012542
;; Publication No. US20030044851A1
;; GENERAL INFORMATION:
;; APPLICANT: Ruben et al.
;; TITLE OF INVENTION: 94 Human Secreted Proteins
;; FILE REFERENCE: P2029P1
;; CURRENT APPLICATION NUMBER: US/10/012,542
;; CURRENT FILING DATE: 2001-12-12
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
;; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
```

;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
 ;; NUMBER OF SEQ ID NOS: 532
 ;; SOFTWARE: Patentln Ver. 2.0
 ;; SEQ ID NO 351
 ;; LENGTH: 101
 ;; TYPE: PRT
 ;; ORGANISM: Homo sapiens
 US-10-012-542-351

Query Match 71.7%; Score 33; DB 9; Length 101;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 DRATCL 8
 |||||
 Db 33 DRATCL 38

RESULT 6
 US-09-821-255-2
 ; Sequence 2, Application US/09821255
 ; Patent No. US20020081293A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Michael S.C. Sun
 ; APPLICANT: Bill N.C. Sun
 ; APPLICANT: Cecily R.Y. Sun
 ; TITLE OF INVENTION: Inhibitors of Complement Activation
 ; FILE REFERENCE: 98-2A
 ; CURRENT APPLICATION NUMBER: US/09/821,255
 ; PRIOR FILING DATE: 2001-03-29
 ; PRIOR APPLICATION NUMBER: 60/075,328
 ; PRIOR FILING DATE: 1998-02-20
 ; PRIOR APPLICATION NUMBER: 09/253,689
 ; PRIOR FILING DATE: 1999-02-20
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 228
 ; TYPE: PRT
 ; ORGANISM: human
 US-09-821-255-2

Query Match 71.7%; Score 33; DB 10; Length 228;
 Best Local Similarity 66.7%; Pred. No. 69;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRATCLR 9
 ::|||
 Db 148 VLDRAATCLR 156

RESULT 7
 US-10-001-843-188
 ; Sequence 188, Application US/10001843
 ; Patent No. US20020132255A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Salceda, Susana
 ; APPLICANT: Macina, Roberto
 ; APPLICANT: Recipon, Herve
 ; APPLICANT: Cafferey, Robert
 ; APPLICANT: Sun, Yongming
 ; APPLICANT: Liu, Chenghua
 ; APPLICANT: Turner, Leah
 ; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and Pr
 ; FILE REFERENCE: DE-0267
 ; CURRENT APPLICATION NUMBER: US/10/001,843
 ; PRIOR FILING DATE: 2001-11-20
 ; PRIOR APPLICATION NUMBER: 60/249,992
 ; PRIOR FILING DATE: 2000-11-20
 ; NUMBER OF SEQ ID NOS: 218
 ; SOFTWARE: Patentln version 3.1
 ; SEQ ID NO 188
 ; LENGTH: 44

;; TYPE: PRT
 ;; ORGANISM: Homo sapien
 US-10-001-843-188

Query Match 69.6%; Score 32; DB 12; Length 44;
 Best Local Similarity 75.0%; Pred. No. 20;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVDRATCL 8
 |||||
 Db 7 LVERAVCL 14

RESULT 8
 US-09-864-761-35244
 ; Sequence 35244, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
 ; FILE REFERENCE: Aecmica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; PRIOR FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 35244
 ; LENGTH: 83
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AL033378.6
 ; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
OTHER INFORMATION: SWISSPROT HIT: P37743, EVALU6 6.90e-01
OTHER INFORMATION: EST_HUMAN HIT: A0141284.1, EVALU6 5.00e-31
US-09-864-761-35244

Query Match 69.6% Score 32; DB 10; Length 83;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 VDRATCLR 9
1 1 1 1 1 1
DB 10 VDRSTCLR 17

RESULT 9
US-10-117-323-38
Sequence 38, Application US/10117323
Publication No. US20030054993A1
GENERAL INFORMATION:
APPLICANT: Rancourt, Derrick E.
APPLICANT: Rancourt, Susan L.
APPLICANT: O'Sullivan, Colleen M.
TITLE OF INVENTION: Implantation Serine Proteinases
FILE REFERENCE: 033337-005
CURRENT APPLICATION NUMBER: US/10/117,323
CURRENT FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: US 60/281,724
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: US 60/294,736
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: US 60/350,962
PRIOR FILING DATE: 2002-01-25
NUMBER OF SEQ ID NOS: 41
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 38
LENGTH: 252
TYPE: PRT
ORGANISM: Mouse
US-10-117-323-38

Query Match 69.6% Score 32; DB 9; Length 252;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LVDRATCLR 9
1 1 1 1 1 1
DB 164 VDRATCSR 172

RESULT 10
US-09-753-436-120
Sequence 120, Application US/09753436
Patent No. US20010029293A1
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Vazeux, Rosemary
TITLE OF INVENTION: ICAM-Related Materials and Methods
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/753.436
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/382,289
FILING DATE:
APPLICATION NUMBER: US 08/487,113
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,754
FILING DATE: 05-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Joseph A., Jr.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 33282
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3836
INFORMATION FOR SEQ ID NO: 120:
SEQUENCE CHARACTERISTICS:
LENGTH: 434 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-753-436-120

Query Match 69.6% Score 32; DB 10; Length 434;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 VDRATC 7
1 1 1 1 1 1
DB 292 IDRATC 297

RESULT 11
US-10-047-542-65
Sequence 65, Application US/10047542
Patent No. US20020168367A1
GENERAL INFORMATION:
APPLICANT: LARRICK, JAMES W.
APPLICANT: WYCOFF, KETH L.
TITLE OF INVENTION: NOVEL IMMUNOADHESINS FOR TREATING AND PREVENTING VIRAL
FILE REFERENCE: 030905.0004.C1P1
CURRENT APPLICATION NUMBER: US/10/047,542
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: PCT/US01/13932
PRIOR FILING DATE: 2001-04-28
PRIOR APPLICATION NUMBER: 60/200,298
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 101

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 65
LENGTH: 547
TYPE: PRT
ORGANISM: Homo sapiens
US-10-047-542-65

Query Match 69.6% Score 32; DB 9; Length 547;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDRATC 7
:|||||
DB 402 IDRATC 407

RESULT 12
US-09-753-436-1
Sequence 1, Application US/09753436
Patent No. US20010029293A1
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
TITLE OF INVENTION: ICAM-Related Materials and Methods
NUMBER OF SEQUENCES: 120
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Garstein, Murray & Borum
SCREEN: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/753,436
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/382,289
FILING DATE:
APPLICATION NUMBER: US 08/487,113
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,754
FILING DATE: 05-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,852
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/009,266
FILING DATE: 22-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/894,061
FILING DATE: 05-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/889,724
FILING DATE: 26-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/827,689
FILING DATE: 27-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Joseph A., Jr.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 33282
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 547 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 30..547
US-09-753-436-1

Query Match 69.6% Score 32; DB 10; Length 547;
Best Local Similarity 83.3%; Pred. No. 2.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 VDRATC 7
:|||||
DB 402 IDRATC 407

RESULT 13
US-09-919-172-29
Sequence 29, Application US/09919172
Patent No. US20020119463A1
GENERAL INFORMATION:
APPLICANT: Paris, Mary
TITLE OF INVENTION: PROSTATE CANCER MARKERS
FILE REFERENCE: PA-0036 US
CURRENT APPLICATION NUMBER: US/09/919,172
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/222,469
PRIOR FILING DATE: 2000-07-28
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PERL Program
SEQ ID NO: 29
LENGTH: 2125
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20020119463A1 3774181CD1
US-09-919-172-29

Query Match 69.6% Score 32; DB 10; Length 2125;
Best Local Similarity 77.8%; Pred. No. 1e+03;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LVDRATC 9
:|||||
DB 1946 LVDRITR 1954

RESULT 14
US-10-012-542-362
Sequence 362, Application US/10012542
Publication No. US20030044851A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: 94 Human Secreted Proteins
FILE REFERENCE: P2029P1
CURRENT APPLICATION NUMBER: US/10/012,542
CURRENT FILING DATE: 2001-12-12
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112

PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
NUMBER OF SEQ ID NOS: 532
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 362
LENGTH: 24
TYPE: PRT
ORGANISM: Homo sapiens
US-10-012-542-362

Query Match 67.48; Score 31; DB 9; Length 24;
Best Local Similarity 75.0%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LVDRATCL 8
1 111111
DB 12 LODRASCL 19

RESULT 15
US-09-870-162A-13
Sequence 13; Application US/09870162A
Patent No. US20020042118A1
GENERAL INFORMATION:
APPLICANT: Brein19, Sabine
APPLICANT: Fuchs, Georg
TITLE OF INVENTION: Phenol Induced Proteins of Thauera aromatica
FILE REFERENCE: BC1006 US DIV
CURRENT APPLICATION NUMBER: US/09/870,162A
CURRENT FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: 09/516914
PRIOR FILING DATE: 2000-03-01
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Microsoft Office 97
SEQ ID NO: 13
LENGTH: 223
TYPE: PRT
ORGANISM: Thauera aromatica
US-09-870-162A-13

Query Match 67.48; Score 31; DB 10; Length 223;
Best Local Similarity 66.78; Pred. No. 16+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LVDRATCL 9
1 111111
DB 46 LITRACCL 54

Search completed: May 6, 2003, 15:04:52
Job time : 19.3333 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 9, 2003, 16:12:15 : Search time 16.9412 Seconds
(without alignments)
153.214 Million cell updates/sec

Title: US-09-851-422a-8

Perfect score: 128
Sequence: 1 GFATLCTKVLDFGIDKLIQLEDKXX 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126	98.4	98	1 S25283	amoebapore A precu
2	87	68.0	96	1 S49144	amoebapore B precu
3	76	59.4	101	1 S49145	amoebapore C precu
4	54	42.2	416	2 S73321	serine/threonine-s
5	54	42.2	421	2 T05032	protein kinase hom
6	53	41.4	835	1 W2B51	gene 51 protein -
7	52	40.6	396	1 D69053	probable iron-sulf
8	51	39.8	520	2 AP2370	serine/threonine k
9	50	39.1	56	2 P10196	protein-tyrosine k
10	50	39.1	213	2 F87669	conserved hypotet
11	50	39.1	499	1 A40092	protein-tyrosine k
12	49	38.3	309	2 S67196	probable membrane
13	49	38.3	507	1 A39939	protein-tyrosine k
14	48.5	37.9	559	2 H71983	DNA primase - HeLa
15	48	37.5	425	2 A99939	proton/sodium-glut
16	48	37.5	481	2 S74590	hypothetical prote
17	48	37.5	118	2 T29915	hypothetical prote
18	47.5	37.1	1274	2 T02636	D1 protein homolog
19	47.5	37.1	559	2 D64521	DNA primase - HeLa
20	47.5	37.1	885	2 D86151	F2248.8 protein -
21	47	36.7	131	2 E71959	hypothetical prote
22	47	36.7	183	2 G95979	conserved hypotet
23	47	36.7	290	2 G36789	hypothetical prote
24	47	36.7	317	2 T45985	hypothetical prote
25	47	36.7	984	2 T48216	hypothetical prote
26	46.5	36.3	859	2 C87502	ATP-dependent RNA
27	46	35.9	131	2 A64550	hypothetical prote
28	46	35.9	215	2 A81410	hypothetical prote
29	46	35.9	539	2 A64441	O-sialoglycoprotei

30	46	35.9	738	1 S51380	protein kinase STE
31	46	35.9	828	1 T00534	S-receptor kinase
32	46	35.9	984	1 A34076	protein-tyrosine k
33	46	35.9	1098	2 T41493	hypothetical prote
34	45.5	35.5	660	2 S73597	hypothetical prote
35	45.5	35.5	1484	2 T29275	lactococcal transp
36	45	35.2	384	2 D97256	probable UDP-gluc
37	45	35.2	445	2 C70076	hypothetical prote
38	45	35.2	485	2 B65072	probable transport
39	45	35.2	485	2 C91098	probable transport
40	45	35.2	485	2 G85943	probable transport
41	45	35.2	505	2 I37206	protein-tyrosine k
42	45	35.2	552	2 C87259	lysyl-tRNA synthe
43	45	35.2	657	2 T48228	probable protein k
44	45	35.2	657	2 G71377	probable transcrip
45	45	35.2	676	2 E96725	hypothetical prote

ALIGNMENTS

RESULT 1
S25283
amoebapore A precursor - Entamoeba histolytica
N:Alternate names: pore-forming protein
C:Species: Entamoeba histolytica
C>Date: 22-Nov-1993 #sequence-revision 09-Mar-1996 #text-change 18-Jun-1999
C:Accession: S25283; A41279
R:Leipe, M.; Tannich, E.; Nickel, R.; van der Goot, G.; Pattus, F.; Horstmann, R.D.; EMO J. 11, 3501-3506, 1992
A>Title: Primary and secondary structure of the pore-forming peptide of pathogenic En
A:Reference number: S25283; PMID:93010939; PMID:1396552
A:Accession: S25283
A:Molecule type: DNA
A:Residues: 1-98 <LE1>
A:Cross-references: EMBL:M83945; NID:g158968; PIDN:AAA2911.1; PID:g158969
A:Note: parts of this sequence, including the amino end of the mature protein, were d
R:Leipe, M.; Ebel, S.; Schoenberger, O.L.; Horstmann, R.D.; Mueller-Berhard, H.J.
Proc. Natl. Acad. Sci. U.S.A. 88, 7659-7663, 1991
A>Title: Pore-forming peptide of pathogenic Entamoeba histolytica.
A:Reference number: A41279; PMID:91352048; PMID:181907
A:Accession: A41279
A:Molecule type: protein
A:Residues: 22-25, 'X', 27-28, 'X', 30-41, 'X', 43-46 <LEM>
C:Superfamily: amoebapore; saposin repeat homology
C:Keywords: antibacterial
F:1-21/Domain: signal sequence #status predicted <SIG>
F:18-98/Domain: saposin repeat homology <SAP>
F:22-98/Product: amoebapore A #status experimental <MAT>
F:26-98, 29-92, 56-67/Disulfide bonds: #status predicted

Query Match 98.4% Score 126; DB 1; Length 98;
Best local Similarity 100.0%; Pred. No. 1.7e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 61 GFATLCTKVLDFGIDKLIQLEDK 85
1 GFATLCTKVLDFGIDKLIQLEDK 25
|||||
amoebapore B precursor - Entamoeba histolytica
S49144
C:Species: Entamoeba histolytica
C>Date: 16-Feb-1995 #sequence-revision 12-May-1995 #text-change 18-Jun-1999
C:Accession: S61440; S61452; S49144
R:Leipe, M.; Andrae, J.; Nickel, R.; Tannich, E.; Mueller-Berhard, H.J.
Mol. Microbiol. 14, 895-904, 1994
A>Title: Amoebapores, a family of membranolytic peptides from cytoplasmic granules of
A:Reference number: S61439; PMID:9523126; PMID:7715451
A:Accession: S61440
A:Molecule type: DNA
A:Residues: 1-96 <LE2>

A: Cross-references: EMBL:X76904; NID:g509761; PIDN:CAAC4226.1; PID:g509762
A:Accession: S61452
A:Molecule type: protein
A:Residues: 20-63 <LEF>
C:Superfamily: amoebapore; saposin repeat homology
C:Keywords: antibacterial
F:1-19/Domain: signal sequence #status predicted <SIG>
F:16-96/Domain: saposin repeat homology <SAP>
F:20-96/Product: amoebapore B #status predicted <MAT>
F:24-96,27-90,54-65/Disulfide bonds: #status predicted

Query Match 68.0% Score 87; DB 1; Length 96;
Best Local Similarity 58.3%; Pred. No. 9.7e-06;
Matches 14; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Oy 1 GFATLCTKRVDFGDIKLIIQLIED 24
 II::III::I::I::I::I::I::I::
Db 59 GRFGTLCKRIISFGVDLAKLTEN 82

```

RESULT 3
S49145
amoebapore C precursor - Entamoeba histolytica
C:Species: Entamoeba histolytica
C:Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 18-Jun-1999
C:Accession: S61439; S61453; S49145
R:Release: M.; Andrade, J.; Nickel, R.; Tannich, E.; Mueller-Eberhard, H.J.
MOL: Microbiol. 14, 895-904, 1994
A:Title: Amoebapores, a family of membranolytic peptides from cytoplasmic granules of E.
A:Reference number: S61439; MUID:95231296; PMID:7715451
A:Accession: S61439
A:Molecule type: DNA
A:Residues: 1-101 <LE2>
A:Cross-references: EMBL:X76903; NID:9509763; PIDD:CAA54225.1; PID:9509764
A:Accession: S61453
A:Molecule type: protein
A:Residues: 25-69 <LE1>
C:Superfamily: amoebapore; saposin repeat homology
C:Keywords: antibacterial
F:1-24/Domain: signal sequence #status predicted <SIG>
F:21-101/Domain: saposin repeat homology <SAP>
F:25-101/Product: amoebapore A #status experimental <MAT>
F:29-101,32-95,59-70/Disulfide bonds: #status predicted

Query Match          59.4%; Score 76; DB 1; Length 101;
Best Local Similarity 59.1%; Pred. No. 0.00044;
Matches 13; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```

```

Oy      1  GFIATLCRKVLDPFGIDKLIQLT 22
      | : ||||| : ||||| : |
Db      64  GLVETLCRKIVSYGIDKLEIK 85

RESULT 4
S75321
serine/threonine-specific protein kinase sll1575 (EC 2.7.1.-) - Synechocystis sp. (str
N:Alternate names: protein sll1575
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S75321
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
S.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S75321
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-416 <K&N>
A:Cross-references: EMBL:ID90904; GB:AB001339; NID:g1652225; PIDN:BA017235.1; PID:d101799
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

```

```

C:Genetics:
A:Start codon: GTG
C:Keywords: phosphotransferase; serine/threonine-specific protein kinase
Query Match      42.2%   Score 54;   DB 2;   Length 416;
Best Local Similarity 62.5%   Pred. No. 3.5;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Oy      9 KVLDFGIDKLILQLIED 24
      ||:||||| ||:| ||:
Db      62 KILDFGIKLVQAAEE 77

```

RESULT 5
T05032
protein kinase homolog F13C5.120 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence,reviston 23-Apr-1999 #text,change 11-Jun-1999
C:Accession: T05032
R:Bevan, M.; Pohl, T.; Weizneger, T.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Sch
submitted to the Protein Sequence Database, February 1998
A:Reference number: Z15395
A:Accession: T05032
A:Molecule type: DNA
A:Residues: 1-421 <BEV>
A:Cross-references: EMBL:AL021711
A:Experimental source: cultivar Columbia; BAC clone F13C5
C:Genetics:
A:Map position: 4
A:Introns: 94/3; 119/3; 151/3; 180/2; 220/3; 247/2; 283/1; 311/3; 352/2; 389/3
A:Note: F13C5.120

Query Match	42.2%	Score 54	DB 2	Length 421
Best Local Similarity	58.8%	Pred. No. 3.5		
Matches	10	Conservative	3	Mismatches 4; Indels 0; Gaps 0;
QY	9	KVLDGFIKLIQLIEDK	25	
Db	258	KVADGVSRLYVREDK	274	

RESULT 6
WZBES1
gene 51 protein - human herpesvirus 3
C:Species: human herpesvirus 3, varicella-zoster virus
C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 16-Jul-1999
C:Accession: G27344
R:Davidson, A.J.: Scott, J.E.
J. Gen. Virol. 67, 1759-1816, 1986
A:Title: The complete DNA sequence of varicella-zoster virus.
A:Reference number: A27345; MUID:86306657; PMID:3018124
A:Accession: G27344
A:Molecule type: DNA
A:Residues: 1-835 <DAV>
A:Cross-references: EMBL:X04370; NID:959989; PIDN:CAA27933.1; PID:960039
C:Genetics:
A:Gene: 51
C:Superfamily: varicella-zoster virus gene 51 protein
C:Keywords: DNA binding

```

Query Match      41.4%: Score 53; DB 1; Length 835;
Best Local Similarity 43.5%: Pred. No. 10;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY      1 GFIAFLCTKRVLDFFGIDKLIQLE 23
      || || : ||| |::|:
DB      225 GFGRTCTILRDMCIDTLVRYIK 247

RESULT 7
D69053
probable iron-sulfur protein MTH140 [similarity] - Methanobacterium thermoautotrophicum
C:Species: Methanobacterium thermoautotrophicum

```

[illegible]

```

A:Molecule type: mRNA
A:Residues: 1-56 <YUE>
C:Experimental source: lymphocyte cell line
C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH
C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
F:1-56/Domain: protein kinase homology (fragment) <KIN>

Query Match          39.1% Score 50; DB 2; Length 56;
Best Local Similarity 53.3% Pred. No. 1-6;
Matches      8; Conservative    4; Mismatches     3; Indels     0; Gaps     0;

QY       5 TLCTKVLDGFDGLKL 19
         ||| :||| ::|
Db        11 TLCKKIADFGLARII 25

RESULT 10
F87669
Conserved hypothetical protein CC3392 [Imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: F87669
R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
B.; Laub, M.T.; Debey, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Hatt, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A>Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB7249; MUID:21173698; PMID:11259647
A:Accession: F87669
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <STO>
A:Cross-references: GB:AE005673; NID:g13425100; PIDN:AAK25354.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3392

Query Match          39.1% Score 50; DB 2; Length 213;
Best Local Similarity 40.0% Pred. No. 6-7;
Matches     10; Conservative     5; Mismatches    10; Indels     0; Gaps     0;

QY       1 GFATLCTRVLDGFDGLQLIEDK 25
         ||| :||| :||| :|
Db        177 GFSASTTIKRSDPGYSKYIPIWGD E 201

RESULT 11
A40092
Protein-tyrosine kinase (EC 2.7.1.112) b1k [validated] - mouse
C:Species: Mus musculus (house mouse)
C:Date: 16-Jun-2000 #sequence_revision 16-Jun-2000 #text_change 16-Jun-2000
C:Accession: A40092
R:Dymecki, S.M.; Niederhuber, J.E.; Desiderio, S.V.
Science 247, 332-336, 1990
A>Title: Specific expression of a tyrosine kinase gene, b1k, in B lymphoid cells.
A:Reference number: A40092; MUID:90117147; PMID:2404338
A:Accession: A40092
A:Molecule type: mRNA
A:Residues: 1-499 <DYMA>
A:Cross-references: GB:M30903; NID:g202076; PIDN:AAA40453.1; PID:g202077
C:Genetics:
A:Gene: MGI:B1k
A:Cross-references: MGI:88169
A:Map position: 14:28.0
C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH
C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation;
F:59-107/Domain: SH3 homology <SH3>
F:118-214/Domain: SH2 homology <SH2>
F:233-491/Domain: protein kinase homology <KIN>
F:241-249/Region: protein kinase ATP-binding motif
F:2/Modified site: myristylated amino end (gly) (in mature form) *status predicted
F:263/Active site: lys *status predicted

Query Match          39.1% Score 50; DB 1; Length 499;

```

Best Local Similarity 53.3%; Pred. No. 16;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 5 TLCTKVLDFGIDKLI 19
||| : ||| : :|

Db 365 TLCKRIADFGFLARI 379

RESULT 12

S67196 Probable membrane protein YOR292c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein O5608

C:Species: Saccharomyces cerevisiae

C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002

C:Accession: S67196

R:Colpelech, C.; Jauniaux, J.C.; Kordes, E.; Polrey, R.; Pujol, A.; Tobiasch, E.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67194

A:Accession: S67196

A:Molecule type: DNA

A:Residues: 1-309 <C21>

A:Cross-references: EMBL:Z75200; NID:91420647; PID:e252132; PID:91420648; GSPDB:GN00015;

A:Experimental source: strain S288C

C:Genetics:

A:Gene: MIPS:YOR292c

A:Cross-references: SGD:S0005818

A:Map position: 15R

C:Keywords: transmembrane protein

F:53-69/Domain: transmembrane #status predicted <TM2>

F:92-108/Domain: transmembrane #status predicted <TM2>

Query Match 38.3%; Score 49; DB 2; Length 309;

Best Local Similarity 56.2%; Pred. No. 14;

Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 4 ATLCTKVLDFGIDKLI 19
||| : ||| : :|

Db 86 ATLCTNLLFGISDIL 101

RESULT 13

A39939 protein-tyrosine kinase (EC 2.7.1.112) tk1 [similarity] - chicken

N:Alternate names: kinase-related transforming protein (tk1); T-cell surface antigen ass

C:Species: Gallus gallus (chicken)

C:Date: 16-Jun-2000 #sequence_revision 16-Jun-2000 #text_change 16-Jun-2000

C:Accession: A42126; A39939

R:Chow, L.M.; Ratcliffe, M.J.; Velllette, A.

Mol. Cell. Biol. 12, 1226-1233, 1992

A:Title: tk1 is the avian homolog of the mammalian lck tyrosine protein kinase gene.

A:Reference number: A42126; MUID:92186854; PMID:1545804

A:Accession: A42126

A:Molecule type: mRNA

A:Residues: 1-88 <CHO>

A:Cross-references: GB:M85043

A:Experimental source: thymus, spleen

A:Note: sequence extracted from NCBI backbone (NCBIN:88831, NCBI:88833)

R:Streibhardt, K.; Mullins, J.I.; Bruck, C.; Ruebsamen-Waismann, H.

Proc. Natl. Acad. Sci. U.S.A. 84, 8778-8782, 1987

A:Title: Additional member of the protein-tyrosine kinase family: the src-and lck-relate

A:Reference number: A39939; MUID:88097370; PMID:3321053

A:Accession: A39939

A:Molecule type: mRNA

A:Residues: 52-507 <STR>

A:Cross-references: GB:J03579; NID:9212712; PIDN:AAA49081.1; PID:9212713

C:Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h

C:Keywords: ATP; autophosphorylation; blocked amino end; lipoprotein; myristylation; phd

F:125-222/Domain: SH3 homology <SH3>

F:241-499/Domain: SH2 homology <SH2>

F:241-499/Domain: protein kinase homology <KIN>

F:249-257/Region: protein kinase ATP-binding motif

F:2/Modified site: myristylated amino end (Gly) (in mature form) #status predicted

F:392,503/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status pred

Query Match 38.3%; Score 49; DB 1; Length 507;

Best Local Similarity 53.3%; Pred. No. 23;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 6 LCTKVLDFGIDKLIQ 20
||| : ||| : :|

Db 374 LCKRIADFGFLARLE 388

RESULT 14

H71983 DNA primase - Helicobacter pylori (strain J99)

C:Species: Helicobacter pylori

A:Variety: strain J99

C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 08-Oct-1999

C:Accession: H71983

R:Alm, R.A.; Ling, L.S.L.; Molt, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.

; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, O.; Taylor, D.E.; Vovis, G.F.

Nature 397, 176-180, 1999

A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p

A:Reference number: A71800; MUID:99120557; PMID:9923682

A:Accession: H71983

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-559 <ARN>

A:Cross-references: GB:AE001441; GB:AE001439; NID:94154511; PIDN:AD05592.1; PID:9415

A:Experimental source: strain J99

C:Genetics:

A:Gene: dnaG

C:Superfamily: DNA primase

Query Match 37.9%; Score 48.5; DB 2; Length 559;

Best Local Similarity 47.8%; Pred. No. 31;

Matches 11; Conservative 3; Mismatches 2; Indels 7; Gaps 1;

OY 6 LCTKVLDFGI-----DKLIQL 21
||| : ||| : :|

Db 149 LCTNKIDYGIENKGLNKDKLIQL 171

RESULT 15

A99939 proton/sodium-glutamate symport protein [Imported] - Staphylococcus aureus (strain N3

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: A99939

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O

ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A99758; MUID:21311952; PMID:11418146

A:Accession: A99939

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-425 <KUR>

A:Cross-references: GB:BA000018; PID:913702333; PIDN:BA043474.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: gltT

C:Superfamily: C4-dicarboxylate carrier protein

Query Match 37.5%; Score 48; DB 2; Length 425;

Best Local Similarity 34.8%; Pred. No. 27;

Matches 8; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 1 GFATLCTKVLDFGIDKLIQLIE 23
| | : | : | : | : | : |

Db 208 GVFATCTTITTFGASALLPLIK 230

Search completed: May 9, 2003, 16:18:48

Job time : 18.9412 secs

THIS PAGE BLANK (0870)